

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2001, 10:04:44 ; Search time 4254.07 Seconds  
(without alignments)  
4400.663 Million cell updates/sec

Title: US-09-325-019-1  
Perfect score: 3658  
Sequence: 1 cggacttaccacccagctcca.....aaaaaaaaaaaaaaaaaaaaa 3658

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_om:\*  
4: gb\_ov:\*  
5: gb\_ph:\*  
6: gb\_pl1:\*  
7: gb\_pl2:\*  
8: gb\_pr1:\*  
9: gb\_pr2:\*  
10: gb\_pr3:\*  
11: em\_fun:\*  
12: em\_hum1:\*  
13: em\_hum2:\*  
14: em\_in:\*  
15: em\_om:\*  
16: em\_or:\*  
17: em\_ov:\*  
18: em\_pat:\*  
19: em\_ph:\*  
20: em\_pl:\*  
21: em\_ro:\*  
22: em\_sts:\*  
23: em\_sy:\*  
24: em\_un:\*  
25: em\_vl:\*  
26: gb\_hcg1:\*  
27: gb\_hcg2:\*  
28: gb\_in1:\*  
29: gb\_in2:\*  
30: em\_ba1:\*  
31: em\_ba2:\*  
32: em\_hum3:\*  
33: em\_hum4:\*  
34: gb\_pr4:\*  
35: gb\_hcg3:\*  
36: gb\_hcg4:\*  
37: gb\_hcg5:\*  
38: gb\_hcg6:\*  
39: gb\_hcg7:\*  
40: em\_hcg1:\*  
41: em\_hcg2:\*  
42: em\_hcg3:\*  
43: em\_hum5:\*

44: gb\_pl3:\*  
45: gb\_pr5:\*  
46: gb\_hcg8:\*  
47: gb\_hcg9:\*  
48: gb\_hcg10:\*  
49: gb\_hcg11:\*  
50: gb\_hcg12:\*  
51: gb\_hcg13:\*  
52: gb\_hcg14:\*  
53: gb\_in3:\*  
54: gb\_hcg15:\*  
55: gb\_hcg16:\*  
56: gb\_hcg17:\*  
57: em\_hcg4:\*  
58: em\_hcg5:\*  
59: em\_hcg6:\*  
60: em\_hcg7:\*  
61: em\_hum6:\*  
62: gb\_hcg18:\*  
63: gb\_hcg19:\*  
64: gb\_ba3:\*  
65: em\_hcg8:\*  
66: em\_hcg9:\*  
67: em\_hcg10:\*  
68: gb\_pr6:\*  
69: gb\_pr7:\*  
70: gb\_hcg20:\*  
71: gb\_hcg21:\*  
72: gb\_hcg22:\*  
73: gb\_hcg23:\*  
74: gb\_ro:\*  
75: gb\_sts1:\*  
76: gb\_sts2:\*  
77: gb\_sy:\*  
78: gb\_un:\*  
79: gb\_vl1:\*  
80: gb\_vl2:\*  
81: gb\_pat1:\*  
82: gb\_pat2:\*  
83: em\_hcg0:\*  
84: gb\_hcg24:\*  
85: gb\_pr8:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2929.4	80.1	176767	63 AF192304	AF192304 Homo sapi
2	2613.4	71.4	2830	10 AF100779	AF100779 Homo sapi
3	920.8	25.2	5020	74 AB004873	AB004873 Mus muscu
4	815.4	22.3	1766	74 AF100777	AF100777 Mus muscu
5	585.4	16.0	138446	35 AC011160	AC011160 Homo sapi
6	422.8	11.6	69016	47 AC023088	AC023088 Homo sapi
7	243.4	6.7	2075	81 AR018957	AR018957 Sequence
8	243.4	6.7	2075	81 I11636	I11636 Sequence 1
9	243.4	6.7	2075	81 I32210	I32210 Sequence 1
10	243.4	6.7	2075	85 H0MCONGRO	H0MCONGRO Human conne
11	243.4	6.7	2312	69 HSCOTGF	HSCOTGF H. sapiens m
12	229.6	6.3	138446	35 AC011160	AC011160 Homo sapi
13	225.4	6.2	2338	74 AB023068	AB023068 Rattus no
14	225.4	6.2	2345	74 AF120275	AF120275 Rattus no
15	223.4	6.1	1496	3 S083916	S083916 Sus scrofa
16	220.4	6.0	2330	74 M08263	M08263 Sus scrofa
17	219	6.0	1953	4 X043524	X043524 Xenopus lae
18	218.8	6.0	1585	4 A0271167	A0271167 Notophtha
19	218.8	6.0	2267	74 M07042	M07042 Mouse FISP-
20	217	5.9	1598	3 U70060	U70060 Sus scrofa
21	207.8	5.7	2330	3 BTA000137	BTA000137 Bos tauru

22	1.99	5.4	1900	69	HSU14750
23	1.98	5.4	1212	10	AF100781
24	196.4	5.4	1973	69	HSNOVH
25	189.2	5.2	1065	74	MMNOV
26	187.6	5.1	1064	74	MMNOV
27	187.6	5.1	2380	74	MMNOV
28	180.4	4.9	1068	10	AF143679
29	179	4.9	1305	74	AF171936
30	173.4	4.7	1927	74	CJUI3063
31	169.4	4.6	2078	4	XIU37063
32	167	4.6	1975	81	A28405
33	167	4.6	1975	81	A28444
34	167	4.6	1806	4	CGNOVNRNA
35	135.2	3.7	1975	4	CHCKCR
36	132	3.6	684	81	A28405
37	129.6	3.5	1283	81	AF083500
38	128.6	3.5	1303	10	AF074604
39	129.6	3.5	1427	10	AF070080
40	126.4	3.5	1871	74	ABO15877
41	126.4	3.5	1987	74	AF218568
42	124.4	3.4	2018	74	MUSCRR61A
43	111.2	3.0	1146	81	EL3814
44	111.2	3.0	1887	81	AF031385
45	111.2	3.0	1935	10	AF003594

## ALIGNMENTS

```

FEATURES
  source
    * 151897 151996: gap of unknown length
    * 151997 176767: contig of 24771 bp in length.
      location/Qualifiers
        1..176767
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="8 map 8q24"
        /clone="CIC-458A3"
        1..51878
        /note="assembly_fragment
        clone_end:17
        vector_side:left"
        15197..176767
        /note="assembly_fragment
        clone_end:SP6
        vector_side:right"
BASE COUNT  47667 a 42712 c 40473 g 45711 t 204 others
ORIGIN
Query Match      80.1%  score 2929.4;  DB 63;  Length 176767;
Best Local Similarity 199.6%  Pred. NO. 0;
Matches 2936;  Conservative. 0;  Mismatches 11;  Indels 0;  Gaps 0;

```

[illegible]

Query Match	Similarity	Score	DB	Length
Best Local Similarity	80.1%	2029.4	DB 63	176767
Matches	2936	Consecutive	0	Mismatches 11; Indels 0; Gaps 0

  

Query	Match	Similarity	Score	DB	Length
QY	638	tacacatcaatgaagcaaggaagatgtctgtgtgtatccagcagaggaatccatga	757		
Db	65994	TCCCTTCCCTTCAGCAGGAGGAAGAACTGCTGCTGTACCGACGAGAGCATTCATGA	66033		
QY	758	cttcaacatttgaggctctgcatatgacacgcctccataaccgaagtactgaggatttg	817		
Db	66054	CTTCACACTCCGGGCTGCATCAGCACAGCCCTCATCAACCCAAAGTATGGAGATTGG	66113		
QY	818	catggaacaatagtgctgcatcccccacaagctcaagatacagacgtgtccctccagtg	877		
Db	66114	CATGCAATATAGTCTCTCATTCCTCTACAGCTAAGACTATGAGAGTGTCTTCCAGTG	66173		
QY	878	tccatgaaggcttggtctctcccgccaggtccatgatgaattaaatgcctgtcttgtaaact	937		
Db	66174	TCCATATGGGCTTGGCTCTCTCCCGCAGAGTCTTATGAGATTATGCTCTGTGTAACCT	66233		
QY	938	gagcgtgtggaatcccaatgatcatctttgtctacttggatctctaccccgacttccaga	997		
Db	66234	GAGCGTAGGAATCCCAATGAAATCTTTGGCTACTTGGAAATCTACCCGTCTTCGAA	66293		
QY	998	aattgccaacatgaggaagcacaacatcttggtgctctggagacaaacaaatgctgtgaaag	1057		
Db	66294	AATTGCCAATGAGGACAGCACAATCTTGGCTTGGGAGACTAACCATAGCCTGTGAAG	66353		
QY	1058	cagtcagcccttatgvgccaatactttccaacaatgagccttagtaaccctgactgtgac	1117		
Db	66354	CAGTCAGCCCTTATGGCCCAATCAATCTTCCACCAATGAGCCTTAGTACCCTATCTGGAC	66413		
QY	1118	ccttggtcctgcacattctctgtctctaaacatccaagaagccgtgagtgtgtctcagggc	1177		
Db	66414	CCTTGCCCTGCATTTCTGTGCTCTTAACCAATCAAAATGACCCCTGATGCTCTCAGGCC	66473		
QY	1178	catgtcatgaattctctccttgataatcatcttcagacatcctctaaagaanaaagtccgtgc	1237		
Db	66474	CATGCTATGAGTTCCTCTCTTATATCATATCAGCATCTACTATAAAGAAAATGCTGTGC	66533		
QY	1238	tctagctgtcttgcgaatacaccacaagcctgtatccagccttccaaagtcatgaagtctt	1297		
Db	66534	TCTAGCTTCTTGAGTACACCCCAAGCCGTGATCCAGCCTTCCAAAGTCACTGAAATGCTCT	66593		
QY	1298	gtctgattcttgcataatccccaagaataggaatcaggtagacttttaataatacctaatt	1357		
Db	66594	GCTGGATCTTGGCTAAATCCCAAGAAATGGAATCAGGTAGACTTTAAATCATCAATTT	66653		
QY	1358	cttctcttagatgcacaacacacagaagcctcttggtatccatcagaatgaatgaatatt	1417		
Db	66654	CTTCTTTAGATGCCAACAACACAGACTCTTTTGGGTCCATTACGATGATATAGATGGAATTT	66713		
QY	1418	ggaacaacatgaataatctattatttggagcctgcacaaagagtaactgtaaatggtaactct	1477		

.Db	66714	GGACAAATAGAAATATCTATTATTGGAGCCTGCCAAGAGTACTGTAAAGGTAAATCT	66773
Qy	1478	gaagtcacgcgcacacaaactactcgtatccaaatgatgatgacactcaaggtcacaa	1537
Db	66774	GACGTCACGCACCAAAACTATCCGATTCCAAATATATGACACTCAAGTATCAAA	66833
Qy	1358	catttgcgaagtgaatgaatagtgtcctaatttgatttctaattgaagagtgatcca	1597
Db	66834	CATTGGCCAAAGAGTGAATGATGTTGCTTAATTTGATTTTAAAGGAAAGTGTATCCA	668933
Qy	1358	ttaactcgggaattgttaggttaagtttctcttccaccctacactgtgaaaggttaacga	1657
Db	66894	TTAACTCTGGCATTTGTTAAGGTTAAGTTTCCTTAACCTTACACTGTGAAGGGATACGA	66953
Qy	1658	ttgaagttgtccacgtcgaagaataatttgaataacattccgttgtatgggaanagccc	1717
Db	66954	TTAGTTGTGTCCCACTCAGAAATTAATAATTGTATAACATTCCTGTTGATGGGAAAGCCC	67013
Qy	1718	ccaagttaaactccacagacagaggaaggtcaagccattccagaaggaaccaatgactct	1777
Db	67014	CCAGTTATTAATCHCCAGAGACAGAGGAAAGTGACGCCGTTTCAGAAGGACCAATTGACTCT	670733
Qy	1778	caacatgaatacagctgcctgactcgtgcagggcctttgggcagttggccaagctcttcctga	1837
Db	67074	CACACTGATCAGCTGCTGACACTGACGAGGGCCTTTGGCAGTTGGCCAGGACTCTTCCTTANA	671333
Qy	1838	tcttcctccctgtccctgtcgtttgggtgttcaatgaatattgaagagctctgnaatgcccgt	1897
Db	67134	TCTTTCTCCCTTGCTCTGCTGGGGTTCAATAGGAATGGTGAAGCCTCTGAGCTGGCCTGT	671933
Qy	1898	ctgagccctctgaagatggtgtccctctggaacactcctactcttacaagagcctctgaagacc	1957
Db	67194	CTGGCCCTTGAGAGTGGTGGCCCTGGAAACACTCCTACTCTTACAGAGGCTTGAAGAAC	672533
Qy	1958	caagctgaagacacatgcagacccactgtgaatgaacagaacaggttcaggttaaggggtgtg	2017
Db	67254	CAGCTGCAGACCATCCACAGACCCACTGGAATATCACCAAGCAGGTTCAAGTAAAGGCTGTGG	673133
Qy	2018	gtcaaaccaagaagaatggtgtccctctgtgtagacagccttggtgtgaacctctagaagctgaagc	2077
Db	67314	GTCAAACCAAGAGTGGGTGCCCTTGTTGTAAGACCTGGGGTATACCTCTTAGACCTTGAGGC	673733
Qy	2078	tgtggagactccaggggcccccggtgtctcaggaacacatcattgtcagaagactcaattccacag	2137
Db	67374	TGTGGAGCTCCAGGGGGCCCCGGTGTTCAGGACACTATTTGCAAGACTCATTTCCACAG	674333
Qy	2138	cccttcgltctgtcgtacccaatgtgcagtlttctgtgttaggaatgtgaggtttaccagtt	2197
Db	67434	CCTTTCGATCTGCTGACCAAAATGGCCAGATTTCTGTGTGAGAAAGTGGAGTTTACCGGTT	674933
Qy	2198	gtttgaaacagaaataaactataaagaagtttaagctctgaagaggttgaagcttaaaagg	2257
Db	67494	GTTTGAACACAAATATGCTTAATAAAGGTTTAAAGCTGAAGAGGTTGAACCTTAAAGG	675533
Qy	2258	aaaaggtgttgttaatgaatacagagctatattattgttaattgaattgagaataataataat	2317
Db	67554	AAAAGGTTGTGTATATGAATATCAGGGCATATTATTATGTATGGAATAATATATATTT	676133
Qy	2318	tactgttggaattcttttaattlaagggcctttctgtgcgaagacattgctcagtgctt	2377
Db	67614	TACTGTTAGAAATCTTTTATTTATTTAGGGCCCTTTCTGTGCACAGCATGTGCTCAGTGCCTTT	676733
Qy	2378	gcattatattagtcactgaatcattcagacaaatgttggaaggttcccatattattctctg	2437
Db	67674	GCAATATTAATGCTCACTGCAATTTTACGCAAAATGTTGAAATGTTCCCATTAATTAATTTCTG	677333
Qy	2438	tctctacaaatgttgaacaggaagctcatagaggttagaanaactcaacagatccaacag	2497
Db	67734	TTCTTACAAATGTGAAGGAAAGCTCATAGAGTGAAGAAATTCATCAACAGATCAACCCAG	677933
Qy	2498	ttgtgtactgtggaagtttagaattcaagatcgaaattggaacttcttaaaccatatt	2557

Dd	67794	TTGTGACTGGGAAGTTAGATTCAAGTCAAGTAATTGGACTGTCTTTATACCCATATT	67855
Qy	2358	tcceccctgttctttagagcttcacaaagtgtgcagaataggaanaacattgcataaattgct	2617
Dd	67854	TCGCCCTGTGTTTAGACTCTCCAAATGTGCAGAAATAGGAAACATTGCATTAATAGGCT	67913
Qy	2618	tgaatttttaattgcaattttccctcttaagtcttcttagctcccttccaagaagcgag	2677
Dd	67914	TGATTTTTTATGTGTCATTTTCCCTCTTAATAGCTTTTAGCTTCCTTTCAAAAGCGAG	67973
Qy	2678	aatttcgatcttcttgtaatttagaggtgcgtctaaagatcccaaaataataggagacacaa	2737
Dd	67974	AATATCGATTTTGTGAATTTTAGTGCTTAAAGCATCCAAATTAATGCGACACACAA	68033
Qy	2738	aatccaggaatccccccttagcttaattccctctcttcccatgsgaacagcctcatcac	2797
Dd	68034	AATCCAGGAATCCCTGTAGTTATTCCTCTTTCCATGTGGAACCAAGCTCATACAC	68093
Qy	2798	attcaaaagatgatctgtcttaeccaaatgctgcataatgaagtgtgtagttaattca	2857
Dd	68094	ATTAAAGATGATTCGTGTTTACCACATCGCATATTGAATGTTGTGATTAATTCACA	68153
Qy	2858	gggaattctgagagtgtgaaagagatctcctaaagggaaagagcctggaataatcc	2917
Dd	68154	GGGATTTCTGTGCAGTGTGCAGAGATTCCTTAACGGGAAAGAGCTGGGAATACATCC	68213
Qy	2918	tccttaactgttaacctccccaacactagtcacagtcaaggtatacagtgtgtctta	2977
Dd	68214	TCCTTAGCTGATCCTCCCCAAACCTAGTCCAGTGCAGAGGTATACAGTGTGCTCATTTA	68273
Qy	2978	atacttgatgaataacaggaagctgtgcaagtgtctcactatttatcgaagctctctct	3037
Dd	68274	ATACTTGATGAATACAGAGAGCTGTGCATGTGTTCTTAATTCGAAGCTCTCTCT	68333
Qy	3038	tccaaagctacatgaanaatgaatttaaacagtcacaaatttatataagtgtcctaaga	3097
Dd	68334	TCCAAAGCTATATGAATGAATGAATTTAACAGTCAAAATTTTATTAATAGCTTACGA	68393
Qy	3098	aagaagacatttaattctcaagaanaatgatatagtatgatacatabatttgatgc	3157
Dd	68394	AAAGGACATTTAATTATTTCCAAAGAAATGCATATGTAATGTAATATATTGTGATGC	68453
Qy	3158	gtaatgcaagaattctgtlataaagaagaattactccaatgaatgacatctctgtlaagta	3217
Dd	68454	GTAATGCAAGATTTCTGTATTAAGAGAAATTCACATCCATGAATGTCCTCTGTAGTCA	68513
Qy	3218	gtgtgatacatggttagatttctgtgaggtgaaaaacacctgcatactcaaaataacaaggc	3277
Dd	68514	GTGTGAATCATGTTTAGATTTTCTGAGAGTGAAGAAACACCTCCCATCAAAATTACAAAGC	68573
Qy	3278	tgtgataaagcttcactcatatttgtaattctagttgaaacccaagaagcttagttctta	3337
Dd	68574	TGTATACAGCTCATCTCATTTGAAATTCAGTGCAGAACCCAAAGCTAGTCTTACTGA	68633
Qy	3338	atttgcaltcaattttggaactgaaacttagcttcttaaaagatcataggaagctgtgtg	3397
Dd	68634	ATTGCAATCTCAATTTGGGAAACTGAACCTTAGCTTCAAGATCATAGGAAGCTGTGTTG	68693
Qy	3398	gagaaactaaggatattatctgcaaatgtgtgtgaggaagagtgtgcagaataaccagtcgc	3457
Dd	68694	GAGAAAGTAGGATTAATTTCTGGCATGTGGTGCAGGAAGGTGTCAAAATTAACCAATGCG	68753
Qy	3458	catgtgttttagaagaagaaactctctatagcagaagccgagaggaaggtcctaaccaca	3517
Dd	68754	CATTGTTTATAGAAACGAACATACTTAATGCAAGCGCCGAGGGCAAGCTTAGACCA	68813
Qy	3518	tgggttgaaagcattgtgtgctgtatccaataaanaattagaataataataatg	3577
Dd	68814	TGGGTGAAAGCCATGGAAGGAATTTGGATCCAAATGATGAAGCGCTTGAAGTCA	68873
Qy	3578	gaattccctgcaatgtgtgtgctgtatccaataaanaattagaataataataatg	3637
Dd	68874	GAATTCCTCTCAATGTGTGGCTGTATCAATTAATAATTAATGAATTAATATATG	68933



```

1384 ATCATTCACACACTACTCTAAGAAAAATGCCCTCTAGCTGTCTGACACACCA 1443
1262 agcctatcagccttccaaagcactagaagcctcgggactctgctaaatcccaag 1321
1444 AGCTTATCCAGCCTTCCAAAGCTAGAGCTGCTGGATCTTGCCCTAAATCCCAAG 1503
1322 aaatgaatcaggtagactttatatacctaattctctctttagatgccaacccaag 1381
1504 AATGGAATCAGGTAGACTTTTAAATCACTAATTTCTTCTTAAATGCCAAACCAAG 1563
1382 actcttgggtccatccagatgaaatagatgaaatttggaacaatagaataatcatt 1441
1564 ACTCTTGGGTCCATTCAGATGAAATGAGATTTGGAACAATAGAAATCTATTATT 1623
1442 tggagcctccaaagaggtactctgtaatgtaattctggaagtcagccaccaacatcc 1501
1624 TGAGAGCTCCCAAGAGGTACTGTAATGGGTAAATTCAGCTGAGCCACCAAAATATCC 1683
1502 tgatccaatatgtaatgcaacctcaaggtcatcaacaacattgccaagtgaatgaat 1561
1684 TGATTCCAATATGTAATGCACTCAAGGTCAATCAATTAATTCGCAAGTGAATAGT 1743
1562 tgcttaatttgattttaaaggaaagtgtatccatcaacctgggcatgtgtgaggtta 1621
1744 TGCTTAATTTTGAATTTTAAATGAAAGTTGTAATCAATTAACCTGGGCATTTGAGGTTA 1803
1622 agtcttcttcaacctacacgtgtaagggttaagattggttcccaagtaagaata 1681
1804 AGTTTCTCTACCCCTACCTGTAAGGTGTAAGATGAGTTGTCCTCCAGTGAAGAAAT 1863
1682 aaattgataaacattccgtgtgtaaggaaagccccaagttaactcctcagaagaaga 1741
1864 AATTTGATTAACATTCCTCTTGTATGGAAAAACCCCACTTAATCTCCAGAGAAAGG 1923
1742 aaagtcagcccatctcagaagaagccaattgactctcaactgaatcagctgtctgag 1801
1924 AAGGTCAGGCCATTTTGAAGAGCAACCAATGACTCTCACACTGAATCAGCTGCTGACTG 1983
1802 cagggcttgggaggtgtgccaagcctctcctggaatctctcctgtctcctgctgg 1861
1984 CAGGGCTTGGGAGGTGGCAAGCTCTCTGTAATCTCTCCCTGCTCTCTT-6GG 2042
1862 ttcatggaattgtaaggcctctggaagcctgtctgctgctgagagtggtgcctct 1921
2043 TTATATGGAATTTGTAAGGCTCTGACACTGCTGTCTGCCCCCTGAGATGGTGCCTG 2102
1922 gaacactcctctactcttcaagagccttgaagagaccagctgcaagaccatgccaagca 1981
2103 GAACTACTCCTCTACTCTTACAGAGCCTTGAGAGACCCAGCTGCAAGCAATGCCAGACCA 2162
1982 ctgaatgacaaagagaggttcagtaggggtgtgggtcaacaagaagtggtgacct 2041
2163 CTGAATGACCAAGACAGGTTCAGTAGGGGTGTGGGTCAACCAAGAAATGGTGGCTCT 2222
2042 tggtaagcagccttggtgagcctctagaagcctgagagcctgtgagagcctcagagggccctg 2101
2223 TGCTACACAGCTTGGGGTGAGCTCTTAAGCTGAGGCTGTGGGACTCCAGGGGCCCCGCTG 2282
2102 ttaagacacatctatctgcaagagccttcaacacacccctctgctctgctgcaacaaatg 2161
2283 TTACAGACACATCTATTTGACAGACATCTTTCACACGCTTCTGCTGCGACCAATAGG 2242
2162 ccaagcttctgtagaagaatgaggttcaacagctgtttagaagaagaatgacttaa 2221
2343 CCAGTTTCTGCTAGAGAAATGAGGTTTACAGTTGTTAGAAACAGAAATAGACTTAA 2402
2222 taaagctttaaagctgaagaggttgaagcctaaagaagaagaagttgttttaagtaatarc 2281
2403 TAAAGGTTTAAAGCTGAAGAGGTTGAAGCTTAAAGGAAAGGTTGTTGTAATGAATATC 2462
2282 aggcataatcttctattagaagaataataattactgcttagaattcttatttag 2341

```

```

Db 2463 AGGCTATTTATTTATTTAGAGAAATATTAATTTACTGTTAGAAATCTTTATTAG 2522
Oy 2342 ggccttctctgccaagacattgctcctcagtgcttgaatgattagctcactgaatct 2401
Db 2523 GGCCTTTCTGTCGCCAGACATTCCTCAGTGTGCTGATTTAGTTCATCATGAAATCTT 2582
Oy 2402 cagcaaatgttgaagaagttccatattattcttcttctacaaatgtaaacggaagc 2461
Db 2583 CAGCAATATTTGAGAAAGTTCCCATATTTCTTCTTTACAAATGTAAGAAACGGAAC 2642
Oy 2462 tcatagaggtgagaacaaccacagagtcacccagtggtgtaatgaggaagtaagatt 2521
Db 2643 TCATAGAGGTGAGAAACCAACCAAGTCAACCACTGCTGAGTGGAAAGTTAGGATT 2702
Oy 2522 cagatcgaatgtgagctgtcttataaccatatttcccgctgttttagagcttcaa 2581
Db 2703 CAGATGGAATTTGAGCTGCTTTATTAACCATATTTTCCCGCTTTTGAAGCTTCA 2762
Oy 2582 atgtgcagaatagaagaacaattgcaataatgagctgatttttaa 2628
Db 2763 AGGTGCAGATAGAGAAACATTCGCAATAATGCGCTGATTTTATA 2809

RESULT 3
AB004873 5020 bp mRNA ROD 25-FEB-1998
LOCUS Mus musculus.mRNA for ELM1, complete cds.
DEFINITION
ACCESSION AB004873.1.GI:2911143
VERSION ELM1; ELM1; expressed in low-metastatic cells.
KEYWORDS Mus musculus (sub species:C3H, strain:HeJ) melanoma cell line;K-1735 M2 CDNA to mRNA, clone_11b:lambda zap II clone:ELM1 R37.
SOURCE
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 5020)
AUTHORS Hashimoto, Y.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-1997) to the DDBJ/EMBL/GenBank databases.
Yasunobu Hashimoto, National Cancer Center Research Institute,
Biology Division, 1-1, Tsukiji 5-chome, Chuo-ku, Tokyo 104, Japan
(E-mail: yhashimoto@ncc.go.jp, Tel:81-3-3542-2511,
Fax:81-3-3542-0807)
REFERENCE
2 (sites)
AUTHORS Hashimoto, Y., Shindo-Okada, N., Tani, M., Nagamachi, Y., Takeuchi, K.,
Shiroishi, T., Toma, H. and Yokota, J.
TITLE Expression of the ELM1 gene, a novel gene of the CCN (connective
tissue growth factor, Cyr61/Cef10, and neuroblastoma overexpressed
gene) family, suppresses in vivo tumor growth and metastasis of
K-1735 murine melanoma cells
JOURNAL J. Exp. Med. 187 (3), 289-296 (1998)
MEDLINE 98119879.
FEATURES
location/Qualifiers
source 1..5020
organism="Mus musculus"
strain="HeJ"
sub_species="C3H"
db_xref="taxon:10090"
cell_line="K-1735 M2"
cell_type="melanoma"
chromosome="15"
clone="ELM1 R37"
clone_11b="lambda zap II"
gene="ELM1"
gene="ELM1"
174..1277
180..1277
180..1277
/gene="ELM1"
/standard_name="expressed in low-metastatic cells"
/product="ELM1"
/protein_id="BAA24949.1"
/db_xref="GI:2911144"

```

/translation="MRLLPRLAVALRVNRLATLSPPTMTTPALEETTT  
 REF:KMPCEPQSPSPRPVSLITDGECKICAOQJGDNCEALIDPRLGLYCD  
 YGDRPRRYAIGCAOVGVGCVLDVNSFCNCTNCTCIDGTGCTPLCLSP  
 RPRLMCRQPRHVRVFGQCEQVDDDDALDTRFAASGAVREORNCIA  
 YTPMSPGCTTGGLSTRISNPNARCNFESRCLNCPVDIOLHKKAKCIAV  
 YDPERNFTLAGCVSTRYPRKYGVCINDNRCIPIYKSKTISVDEQCEGFGFROV  
 LMNNCFULSCRNPMDITADESPDEFERIAN"

BASE COUNT 1341 a 1173 c 1145 g 1361 t  
 ORIGIN

Query Match 25.2%; Score 920.8; DB 74; Length 5020;  
 Best Local Similarity 60.9%; Pred. No. 4.3e-209;  
 Matches 1938; Conservative 0; Mismatches 1092; Indels 152; Gaps 21;

QY 5 cttaccocccagctccacatgtagagacacatctccaccccaattctgcaagtgcccatg 64  
 DB 272 CTTACACCCCGACACACTAGAGAAAGACTACACGCCCGAATTCTGCAAGTGGCCATG 331  
 QY 65 tgaagtccegcacatcccccacccctgcccgtgggggtacgctcatcacaagatgctg 124  
 DB 332 TGAATGCTGTAAGATATGTGCTCCAGCACTGCGCTCCAGCTGAGCGCTACAGATGAGCTG 391  
 QY 125 tgaagtcctgtaagatgtagctcagacagcttgggagacatgacagagagctgcatctg 184  
 DB 392 TGAATGCTGTAAGATATGTGCTCCAGCACTGCGCTCCAGCTGAGCGCTACAGATGAGCTG 451  
 QY 185 tgaaccccaacccgggctcctactgtactacacagcgagccgcccagagtaacagatag 244  
 DB 452 TGACCCACACCGGGGGCTCTACTGCGATTACAGTGGGGATGCGCCGAGATACCAATAGG 511  
 QY 245 agtgctgacagagtgtagtggtggtgctgctgctgagatgaggtgctgctacacaag 304  
 DB 512 AGTGTGTGACAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 571  
 QY 305 ccaagtcctccagctcactgacagacagacatgacagctgacagagcggtgggtgctg 364  
 DB 572 CGAGTCTCTCAACCCACACTGACAGTACACTGATGATGATGATGATGATGATGATGATG 631  
 QY 365 caacacactgtgctcctcagatgtagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 424  
 DB 632 CACACCGCTGTGCTAAAGCCCAAGGCCCAAGCCCTGTGCTGCTGCTGCTGCTGCTGCTG 691  
 QY 425 gtagcactcctgacacactgctgtgtagcagtggtatgtgtagacagacagacagacag 484  
 DB 692 GAAAGTCCCTGGCCAGTGTGCTGAGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 751  
 QY 485 caagagacagaccccccagacacagagacagacagacagacagacagacagacagacag 544  
 DB 752 CCAAGCTGCACTGTGACACACAGAGCTTTGACGCTGACGCGCCGTGAGACAGCGTA 811  
 QY 545 caagaaactgatacctacacaaagccctgtagcccttctgctcacaacagctgctgctg 604  
 DB 812 TGAAGAACTGATAGCCCTACACTAGTCCCTGAGACCCCTGCTTACCACTGCTGAGCTT 871  
 QY 605 ggtctcactcagatctccaaatgttaaagccagtgctggtcctgtagcaagaagagagc 664  
 DB 872 TATCTCCACTCGGATCTCTAAGCTCAATGCCCGCTGCTGAGACAGAGAGAAAGTGCCT 921  
 QY 665 ctgcaacttgagcagctgtagtgtagacatccatcaatcaagagcagagagagagag 724  
 DB 921 CTGCAACCTGGCCAGCATGATGTGACATCACTAATCAATCAAGAGAGAGAGAGAGAG 981  
 QY 725 tctgagctgtgtacacagcagagacatctgtaactcaacttctggtggtgctgcatcac 784  
 DB 981 CCGTGGCTGTGTACACAGCAG 1041  
 QY 785 acgctcctatcaaacacagactgtgtgagttgcatgagacatagagtgctgcatcccta 844  
 DB 1041 ACGCAGCTTACGAGACCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1101  
 QY 845 caagcttaagactacagagtgctcctcagtgctgtagtggtggtggtggtggtggtggt 904

DB 1112 CAAAGTCAAGACCATATAGTGTGATTTCCAGTGTCCAGAGAGGGGCCAGTTCCTCCGGCA 1171  
 QY 905 ggtccatgataatgagctgctctctgtaaccctgagctgtaggaatcccaatgacatct 964  
 DB 1172 GGTCTATGATTAATGATGCTGCTGTGCAACCTGAGCTGACAGAAATCTCAAGATATCTT 1231  
 QY 965 tgcctgacttggaatctcaccctgactctcagaatctcccaactagagcagcacaactct 1024  
 DB 1232 TGCTGACTTGGAATCTTACCTGACTTCCAGAGATGCTCAATTAAGTGGGTGTGAGCT 1291  
 QY 1025 tgggtcttgaggagac 1084  
 DB 1292 CAGGCAAAAGTTTCAAGCTGCAAGACACACACACACACACACACACACACACACAC 1351  
 QY 1085 tcacaaatgagccctagttac--cctgactgagccctgagccctgagccctgagccctg 1142  
 DB 1352 AGCCTTATTTCACTTCTTCT 1411  
 QY 1143 caatcaaatgagcctgtagtgctgctcagagcacaatgctatgagttctccttgata 1202  
 DB 1412 CAATCGTAAATGGCCAGAGAGAGTGTCTCAGGCTCAGACAAATG--GTTCTCTCTG 1470  
 QY 1203 tcatltagatctacactcaagaagaatgctgctcctgctgctgctgctgctgctgctg 1262  
 DB 1471 ACATTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 1530  
 QY 1263 gctctgacagccttccacagctcactcagagctcctgagctcctgagctcctgagctc 1319  
 DB 1531 GGCACAGCTAGTGTGCTCAGCCTGAGCAAGTGTGCAAGTGTGATGGATGTGTGCA 1590  
 QY 1320 agaaatggaatcagagagcttcaatcaatcaatcaatcaatcaatcaatcaatcaatca 1379  
 DB 1591 AGGAAAGATCAGCTGAGAAACCAAGTATCAAGTATCAAGTATCAAGTATCAAGTATCA 1649  
 QY 1380 agactcttggtgctcactcagatgagatgagatgagatgagatgagatgagatgagatg 1439  
 DB 1650 GGGATCTGGAATCTTCAAGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1709  
 QY 1440 ttgtagccttccag 1499  
 DB 1710 TTTTACCTTGCCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1769  
 QY 1500 cctgattccaataatgtagcactcaagagctcaacacattgccaagtgagtgata 1559  
 DB 1770 TCTTGTTCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1828  
 QY 1560 gttgcttaatttgaatttaagagagagagagagagagagagagagagagagagagag 1609  
 DB 1829 GTTGTCTATCTGTTGTCTATGAAATCATTTAAAAAGAGAGAGAGAGAGAGAGAGAG 1888  
 QY 1610 -----ttgttgaagttagttctctcaccacacacacacacacacacacacacacac 1645  
 DB 1889 AAAAGAAAGAAAGATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 1948  
 QY 1646 gaaagttcaagatgagttgttccagtcagaaataaataattgataaataatttctgtga 1705  
 DB 1949 AAAGGCTAAAGTGGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2008  
 QY 1706 tgggaagagcccccagttat--actcagagagcagagagagagagagagagagagag 1763  
 DB 2009 AGGGAAGTGTGAGCTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTTCAATTT 2068  
 QY 1764 gacaaatgactctcacaactgaaatcagctgtagcagcagagccttggcagcttgagca 1823  
 DB 2069 CCAATACCTGGCAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2126  
 QY 1824 ggtctcctctgagcttctcctgctgctgctgctgctgctgctgctgctgctgctgctg 1883  
 DB 2127 -----AAATCTGGGCT 2174  
 QY 1884 ctgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1943  
 DB 2175 GTAGTGAATCTGTCTGTCTCAAAAAGTAAGATGAAAGA-----TGTTCATG 2227





Db 177 CTTCACCCGACACACTAGAGAAAGACTACAGCCCGAATTCTGCAAGTGGCCATG 236  
 QY 65 tgaagtccgcacatcccccaccccgctgcgcgtggtggttcagctctcaacatgctg 124  
 Db 237 TGAGTGGCCACATCCCACTGCTGCCACTGGGGGTGAGCTTATCAAGATGGCTG 236  
 QY 125 tgaagtctgaagatgctgcacagcgttgggagcaactgcaaggagctgcacatg 184  
 Db 297 TGAATGCTGAATATATGTCCCAAGCTTGGGAGCAATGCACAGAGGCTGCATCTG 356  
 QY 185 tgaagcccccacccgggctctcaactgctacagcagggagacccgggagagcaatagg 244  
 Db 357 TGAACCCACACCGGGGCTCTACTGCGATTACAGTGGGAGTCCCGGAGTACCAATAGG 416  
 QY 245 agtctgtgcacaggtgctgctgtgtgctgcctgcctgagctgggtgtgctacacacg 304  
 Db 417 AGTGTGTGCACAGGTGCTGCTGTGGGCTGTGTCTGTGGATGGCTTACACCAATGG 476  
 QY 305 ccaagtctccagcctaactgcaagtaacaaactgcaactgcaactgcagcggtgtggtg 364  
 Db 477 CGAGTCTCTCAACCAACTGCAAGTGAATGTAATGTAATGTAATGTAATGTAATG 536  
 QY 365 ccaacacactgtgctccagtgatgagcccccgcgtctgtgtgtgtgtgtgtgtgtgt 424  
 Db 537 CACACCGCTGTGTAAAGCCCAAGCCCAAGCCCTGTGTGTGTGTGTGTGTGTGTGT 536  
 QY 425 gacatatacctgacacactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 484  
 Db 597 GAGAGTCCCTGGCCAGT 656  
 QY 485 caaagacccgaccccgtagacacagagccttcagatgctgtgtgtgtgtgtgtgtgtgt 544  
 Db 657 CCAGACTGACTGTGACACACAGAGCCTTTCAGCTGAGCTGAGCTGAGCTGAGCTG 716  
 QY 545 caagaaatgcatgaccccaacacagccctgtgagccctgtgtgtgtgtgtgtgtgtgt 604  
 Db 717 TGAGAACTGTGATGCTTACACTAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 776  
 QY 605 ggtctccacacgagctctcaactgtttaaagccagctgtgtgtgtgtgtgtgtgtgtgt 664  
 Db 777 TATCTCCACTCGGATCTCTAAGCTCAATGCCCGGTGTGTGTGTGTGTGTGTGTGT 836  
 QY 665 ctgcaactgctgcgcacatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 724  
 Db 837 CTGCAACTCTGGGCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 896  
 QY 725 tctggtctgtgtacacagcagagcagcagcagcagcagcagcagcagcagcagcagc 784  
 Db 897 CTTGGCTGT 956  
 QY 785 acgctctac 844  
 Db 957 ACCGACTTACCGACCAAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1016  
 QY 845 caagcttaagacatcagcagctgtctcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 904  
 Db 1017 CAAGTCCAAACACATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1076  
 QY 905 ggtctctatgataatcctgt 964  
 Db 1077 GGTCTTATGATTAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1136  
 QY 965 tgcgtacttgatactcactcactcactcactcactcactcactcactcactcactcactc 1024  
 Db 1137 TGCGGACTTGGAACTTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1196  
 QY 1025 tgggtctgt 1084  
 Db 1197 CAGGCTTAAATCTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1256  
 QY 1085 tcaacaaatgagccttagttag--cctgactgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1142

Db 1257 AGCTTATTTATCTACTCTCTACTCTCTACTCTCTACTCTCTACTCTCTACTCTCT 1316  
 QY 1143 ccattcaaatgagcctgt 1202  
 Db 1317 CATCTGTAATGGCCAGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1375  
 QY 1203 tcaatgaactactactactactactactactactactactactactactactactactact 1262  
 Db 1376 ACATTTACATCATCTTCAAGAAACACATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1435  
 QY 1263 gctgtatcagcctctcagacagcagcagcagcagcagcagcagcagcagcagcagc 1319  
 Db 1436 GGTCCAGCTTATGT 1495  
 QY 1320 agaaatgaaatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1379  
 Db 1496 AGGAAAGATCATGCTGGAAGAACCAATATCATATCAATCTCTCTCTAGATGCCAAGCC 1554  
 QY 1380 agactctgt 1439  
 Db 1555 GGGATCTGGGATCTTTCAGACAGATGATGGATGGGAGACACAGATTAAGCTATTA 1614  
 QY 1440 ttgtgagcctgt 1498  
 Db 1615 TTTTACCCTTTCACAAATGATCTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1674  
 QY 1499 tctgtatccaaatgatactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1558  
 Db 1675 TTCTTGTTCACGATCTGTATATCAAGTCAACCAATCTTTCAGGTGAGGACCAT 1734  
 QY 1559 agtcttaattgt 1577  
 Db 1735 AGTTGTCTATCTGT 1753

RESULT 5  
 AC011160 138446 bp DNA HTG 30-MAR-2000  
 LOCUS  
 DEFINITION Homo sapiens clone RP11-9F9, LOW-PASS SEQUENCE SAMPLING.  
 AC011160.2 GI:7341861  
 AC011160.2 GI:7341861  
 VERSION  
 KEYWORDS HTG; RTGS\_PHASE0.  
 SOURCE  
 ORGANISM human.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE  
 1 (bases 1 to 138446)  
 Biren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens, clone RP11-9F9  
 Unpublished  
 2 (bases 1 to 138446)  
 Biren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckwith, R., Boguski, L., Boulikas, B., Brown, A., Castle, A., Collins, S., Collins, S., Collymore, A., Cooke, P., DeRubeis, K., Dewar, K., Domini, M., Donnell, L., Doyle, M., Galagan, J., Garg, N., Grant, G., Hargis, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J., Lebecky, J., Liu, C., Locke, K., MacDonald, P., Margulis, N., McQuinn, P., McQuinn, A., McKernan, K., McLaughlin, J., Melnick, J., Morrow, J., Naylor, J., Norman, C., O'Connor, T., O'Donnell, P., Petersen, J., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange, T., Stojanovic, N., Subramanian, A., Talamas, J., Teste, S., Tittel, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, M., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 COMMENT  
 All repeats were identified using RepeatMasker:  
 Smit, A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/XK/RepeatMasker.html

Direct Submission  
 Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 920 Charles Street, Cambridge, MA 02141, USA  
 On Mar 30, 2000 this sequence version replaced g1:6006142.  
 All repeats were identified using RepeatMasker:  
 Smit, A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/XK/RepeatMasker.html



Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
Project Information  
Center project name: L3009  
Center clone name: 9\_F\_9

\* NOTE: This record contains 152 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 798: contig of 798 bp in length  
799 898: gap of 100 bp  
899 1669: contig of 771 bp in length  
1670 1769: gap of 100 bp  
1770 2540: contig of 771 bp in length  
2541 2640: gap of 100 bp  
2641 3396: contig of 756 bp in length  
3397 3496: gap of 100 bp  
3497 4286: contig of 790 bp in length  
4287 4386: gap of 100 bp  
4387 5167: contig of 781 bp in length  
5168 5267: gap of 100 bp  
5268 6055: contig of 788 bp in length  
6056 6155: gap of 100 bp  
6156 6940: contig of 785 bp in length  
6941 7040: gap of 100 bp  
7041 7834: contig of 794 bp in length  
7835 7934: gap of 100 bp  
7935 8740: contig of 806 bp in length  
8741 8840: gap of 100 bp  
8841 9619: contig of 779 bp in length  
9620 9719: gap of 100 bp  
9720 10498: contig of 779 bp in length  
10499 10598: gap of 100 bp  
10599 11352: contig of 754 bp in length  
11353 11452: gap of 100 bp  
11453 12245: contig of 793 bp in length  
12246 12345: gap of 100 bp  
12346 13127: contig of 782 bp in length  
13128 13227: gap of 100 bp  
13228 14019: contig of 792 bp in length  
14020 14119: gap of 100 bp  
14120 14915: contig of 796 bp in length  
14916 15015: gap of 100 bp  
15016 15784: contig of 769 bp in length  
15785 15884: gap of 100 bp  
15885 16669: contig of 785 bp in length  
16670 16769: gap of 100 bp  
16770 17539: contig of 770 bp in length  
17540 17639: gap of 100 bp  
17640 18432: contig of 793 bp in length  
18433 18532: gap of 100 bp  
18533 19296: contig of 764 bp in length  
19297 19396: gap of 100 bp  
19397 20173: contig of 777 bp in length  
20174 20273: gap of 100 bp  
20274 21046: contig of 773 bp in length  
21047 21146: gap of 100 bp  
21147 21933: contig of 787 bp in length  
21934 22033: gap of 100 bp  
22034 22823: contig of 790 bp in length  
22824 22923: gap of 100 bp  
22924 23706: contig of 783 bp in length  
23707 23806: gap of 100 bp

23807 24579: contig of 773 bp in length  
24580 24679: gap of 100 bp  
24680 25481: contig of 802 bp in length  
25482 25581: gap of 100 bp  
25582 26374: contig of 793 bp in length  
26375 26474: gap of 100 bp  
26475 27247: contig of 773 bp in length  
27248 27347: gap of 100 bp  
27348 28133: contig of 766 bp in length  
28134 28233: gap of 100 bp  
28234 29016: contig of 783 bp in length  
29017 29116: gap of 100 bp  
29117 29882: contig of 766 bp in length  
29883 29982: gap of 100 bp  
29983 30718: contig of 736 bp in length  
30719 30818: gap of 100 bp  
30819 31594: contig of 776 bp in length  
31595 31694: gap of 100 bp  
31695 32454: contig of 760 bp in length  
32455 32554: gap of 100 bp  
32555 33312: contig of 758 bp in length  
33313 33412: gap of 100 bp  
33413 34176: contig of 764 bp in length  
34177 34276: gap of 100 bp  
34277 35068: contig of 792 bp in length  
35069 35168: gap of 100 bp  
35169 35960: contig of 792 bp in length  
35961 36060: gap of 100 bp  
36061 36833: contig of 773 bp in length  
36834 36933: gap of 100 bp  
36934 37703: contig of 770 bp in length  
37704 37803: gap of 100 bp  
37804 38592: contig of 789 bp in length  
38593 38692: gap of 100 bp  
38693 39480: contig of 786 bp in length  
39481 39580: gap of 100 bp  
39581 40371: contig of 791 bp in length  
40372 40471: gap of 100 bp  
40472 41251: contig of 780 bp in length  
41252 41351: gap of 100 bp  
41352 42138: contig of 787 bp in length  
42139 42238: gap of 100 bp  
42239 43010: contig of 772 bp in length  
43011 43110: gap of 100 bp  
43111 43884: contig of 774 bp in length  
43885 43984: gap of 100 bp  
43985 44768: contig of 784 bp in length  
44769 44868: gap of 100 bp  
44869 45643: contig of 775 bp in length  
45644 45743: gap of 100 bp  
45744 46510: contig of 767 bp in length  
46511 46610: gap of 100 bp  
46611 47388: contig of 778 bp in length  
47389 47488: gap of 100 bp  
47489 48270: contig of 782 bp in length  
48271 48370: gap of 100 bp  
48371 49160: contig of 790 bp in length  
49161 49260: gap of 100 bp  
49261 50035: contig of 775 bp in length  
50036 50135: gap of 100 bp  
50136 50916: contig of 781 bp in length  
50917 51016: gap of 100 bp  
51017 51801: contig of 785 bp in length  
51802 51901: gap of 100 bp  
51902 52690: contig of 789 bp in length  
52691 52790: gap of 100 bp  
52791 53578: contig of 788 bp in length  
53579 53678: gap of 100 bp  
53679 54447: contig of 766 bp in length  
54448 54547: gap of 100 bp  
54548 55343: contig of 796 bp in length  
55344 55443: gap of 100 bp  
55444 56213: contig of 770 bp in length



```

38696      39446: contig of 751 bp in length
39447      39546: gap of 100 bp
39547      40280: contig of 734 bp in length
40281      40380: gap of 100 bp
40381      41191: contig of 811 bp in length
41192      41291: gap of 100 bp
41292      42094: contig of 803 bp in length
42095      42194: gap of 100 bp
42195      42968: contig of 774 bp in length
42969      43068: gap of 100 bp
43069      43860: contig of 792 bp in length
43861      43960: gap of 100 bp
43961      44746: contig of 786 bp in length
44747      44846: gap of 100 bp
44847      45630: contig of 784 bp in length
45631      45730: gap of 100 bp
45731      46524: contig of 794 bp in length
46525      46624: gap of 100 bp
46625      47420: contig of 796 bp in length
47421      47520: gap of 100 bp
47521      48330: contig of 810 bp in length
48331      48430: gap of 100 bp
48431      49224: contig of 794 bp in length
49225      49324: gap of 100 bp
49325      50128: contig of 804 bp in length
50129      50228: gap of 100 bp
50229      51042: contig of 814 bp in length
51043      51142: gap of 100 bp
51143      51925: contig of 783 bp in length
51926      52025: gap of 100 bp
52026      52843: contig of 818 bp in length
52844      52943: gap of 100 bp
52944      53708: contig of 765 bp in length
53709      53808: gap of 100 bp
53809      54587: contig of 779 bp in length
54588      54687: gap of 100 bp
54688      55468: contig of 781 bp in length
55469      55568: gap of 100 bp
55569      56414: contig of 846 bp in length
56415      56514: gap of 100 bp
56515      57335: contig of 821 bp in length
57336      57435: gap of 100 bp
57436      58239: contig of 804 bp in length
58240      58339: gap of 100 bp
58340      59150: contig of 811 bp in length
59151      59250: gap of 100 bp
59251      60045: contig of 795 bp in length
60046      60145: gap of 100 bp
60146      60950: contig of 805 bp in length
60951      61050: gap of 100 bp
61051      61806: contig of 756 bp in length
61807      61906: gap of 100 bp
61907      62688: contig of 782 bp in length
62689      62788: gap of 100 bp
62789      63551: contig of 763 bp in length

```

Db	38920	GTGAGAAAAC	TACACGAGTCA	CCCAAGTGGT	GTGACTGGGAA	AGTTAGATTCA	GATTCGA	38979
QY	2530	aatygactgtc	tcttataaac	ccatatttcc	cccgltttt	agagcttcaaa	tgtgtca	2589
Db	38980	AATTGACATG	TCCTTATAAC	CCCATATTTT	CCCCCGTTT	TAGAGCTTCA	AAATGTCTA	39039
QY	2590	gaattggaaa	aaacattgcaa	taaaigtgt	gatcttttta	gtgcatcttcc	ccctcatag	2649
Db	39040	GAAATGGAAA	CAATTCGAAT	TAATGGCTGAT	TTTTTAATG	TCATTTTCC	CTTAATG	39099
QY	2650	tcttctagct	ccctttccaa	aagacagaga	tatactgatttc	-tgaataatt	agagctt	2708
Db	39100	TCTTCTAGCT	CCTTTTTC	CAAGAGAGAA	ATCTGATTTT	CTTGTAATTT	TAGAGTGCTT	39159
QY	2709	aagcatccaaa	tacatgagaca	cac--aaaa	atccagaa	tccccgt	tagctt--att	2763
Db	39160	AAGGCTTCCA	AAATCCATG	GGGACAC	CCCAAAATCC	CAAGATTC	CCCTGTGCTTAAT	39219
QY	2764	ccctcttccca	tcggaaacc	2783				
Db	39220	CCCTTTTCC	CAATTGGGAC	39239				

	RESULT	7					05-DEC-1996
	AR018957						
	LOCUS	AR018957	2075 bp	DNA	PAT		
	DEFINITION	Sequence 1 from patent US 5783187.					
	ACCESSION	AR018957					
	VERSION	AR018957.1	GI:3974071				
	KEYWORDS	.					
	SOURCE	Unknown.					
	ORGANISM	Unknown.					

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
1 (bases 1 to 2075)	1 (bases 1 to 2075)				
	Grötdorff, G.R. and Bradham, D.M., Jr.	Treatment of cell proliferative disorders using antibodies which bind connective tissue growth factor			
	Patent: US 5783187 A 1 21-JUL-1996;				
	Location/Qualifiers				
	1. .2075				

```

/organism="unknown"
BASE COUNT      491 a      558 c      546 g      480 t
ORIGIN

```

Query Match	6.7%	Score	243.4	DB	81	Length	2075
Best Local Similarity	56.2%	Pred	No.2.4e-47				
Matches	523	Conservative	0	Mismatches	396	Indels	12
						Gaps	3

QY	82	caaccgcgtgcgcgcgttggtgggttcacgcccatacacagatgtgtgagtgtcgtataagttgt	1414
Db	248	CGCCCGCGTCCCGGCGGCGCGGTGAACCTCTGCTGGAGAGGTGGCGGTGCTGCGCCGCGTCT	307
QY	142	gcgcctcagcagcttctggggacaaatctgacagggagctctgcacattctgtgaccccccaaccgggccc	201
Db	308	CGCCCAACAGACTGGCGGCAAGCTGTGCACCGAGCGCGACACCTTGGGACCCGACAAAGGCC	367
QY	202	tctactgtgactacaagcggggagccgcccgggtacgcgaattagagttgtgtgacaaagttgg	261
Db	368	TCTTCTGTGACT---TGCGCTCCCGGCGCAACCGCAAGATGGCGTGTGCACCGCCAAAG	424
QY	262	tccggtgtggtcgtcgctccctgtgagtgggtgtgcgtacaacaaacagcagatcccttcagacta	321
Db	425	ATGGGCTCCCTGCATCTTCGGTGGTACGGGTGACCCACGCGGAGAGATCCTTCCAGAGCA	484
QY	322	actgtaagtatacaatcgcgcacgtgatctgacgcggcggtgtgggtctgtgacaaacacattgt---	378
Db	485	GCTGCAATACCAAGTGCACAGCTGCCTGGACGCGGCGGCGGTGGGCTGCATCCCTCTGTACAGCA	544
QY	379	tccgagtgcgcccccccggtctctcgtgtgcccccaaccggcggtgtgagcatatccttggcc	438
Db	545	TGGAGTTCGTCTGCGCACGCCCTACTTCGCCCTTCCGAGGAGGGTAAAGCTGCCCGGA	604

OY	439	accgtctgagacgaatgggtatgtaggaagcagcgcaagagcgacacgaagacggacaccc	438
Db	605	AATGCTGGAGAGAGTGGGTGTGTACGAGCCCAAGGACCAAAACCGTGGTGGCCCTCC	664
OY	459	gtgaacacagygcccttcgatcgcttggtggtgagtgagggcattgacagga-----act	552
Db	655	TGCGGCGCTTACCGACTGGAAACACAGCTTTGGCCCAAGACCAACTATATTAGAGCCAACT	724
OY	553	gcataagccttacaagaagcccccttgagccccctgtctccacacagctcgcgccctgggtctcca	612
Db	725	GCCCTGGTCCAAACACACAGAGTGTGAGCGCCCTGTTCCAAAGACCTGTGGGATGGCATCTCCA	784
OY	613	ctcgatctccgaatggttaagcccccaagtgcttgagccttgagcaagagagcgccctctgcgaact	672
Db	785	CCCCGATTACCAATGACACACCGCTCTCTGACAGCTTAGAGAAACCAAGCGCCCTGTGCATGG	844
OY	673	tgcgcgcattgagatgtgtagacatccataacatccattaaagcagggaaagagtgctgctg	732
Db	845	TCAAGCCTTTGGGAAGCTGACCTGGAGAGAGAACTATTAGAAAGCGCAAAAGTCATCCGTA	904
OY	733	tgtaccagcagagagcatccatctgaacttcacaattcgcggtgcataagcaagcagctcct	792
Db	905	CTCCCAAAATCTCCAGCGCCTATCAAGTTTGACTTCTTGCTGGCTGCACACAGATGAAGCAT	964
OY	793	atcaaccccaagtactgtgtaggttgcatctgtagcaatagtggtgtgataccccaagagcta	852
Db	965	AACGAGCTAAATTTGTGTGAGATATGTACCGAGCGCCCATGTGTGACACCCCCACAGAAACA	1024
OY	853	agactatcgaatgtcttccttcacagtgctcctgtagtggtccttcctccgcagagttccat	912
Db	1025	CCACCCCTCCGGTGGAGTTCAAAGTGGCCCTGACGGCGAGGTCTATGAAGAAGACATATGT	1084
OY	913	ggatthaatgcctgtctctctgtaacctgagcctgttaggaatcccaatgatcatcttgcgact	972
Db	1085	TCATCAAAAGACCTGTGGCTCGCATTCACAACTCTCCCGAGACAAATGACATCATCTTTGAATCGC	1144
OY	973	tggaaatcctaccctgacctcttcagaatatgc	1003
Db	1145	TGTACTACAGGAGATGTATCGAAGACATATGGC	1175

[illegible]

ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 2075)
TITLE	Grondorff G.R. and Bradham D.M. Jr.
JOURNAL	Connective tissue growth factor(CTGF)
FEATURES	Patent: US 5408040-A 1 18-APR-1995;
source	Location/Qualifiers
	1..2075

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      491 a 558 c 546 g 480 t
ORIGIN

```

Query Match:	6.7%	Score 243.4;	DB 81;	Length 2075;
Best Local Similarity:	56.2%	Pred. No. 2.4e-47;		
Matches 523;	Conservative 0;	Mismatches 396;	Indels 12;	Gaps 3

QY 82 caccgcgacggccgcgtcggtcagccatcacagtgtctgtagtcgtctaagatgt 141  
| | | f l k i l l l l l | | | | | | | | | | | |  
Db 248 cgccgccctacccggcgccggaaccttgcctggacggcctgcccgtccgcctct 307

QY	142	ggcgtcagcagcttgggggaaacacgcgcaggaagcgtgcacatctcttgaccccccacggggcc	201
Db	308	ggcgcgaagcagcttggggagctgtgtgcacccagcgcaccccttcgcagaccgcacaagggcc	367
QY	202	tctacttgactacagcaggggaacccgcggcggtacgcataagagtgctgtgcacaggtgg	261
Db	368	ttctttctgtaact---tcgggctccccggcgaacccgacagatcgggggtgtgacccggcmaag	424
QY	262	tcgggtgtggagctcgtcccttgatgggtgtgcgtctacaacacggcagctctccagccta	321
Db	425	atgggtgctccctgcactcttgcgtgtgtaacgggtgtacccacacggcgagatccttcagagca	484
QY	322	actcagaatataacttgacagcgtgcatacgcgcggcggtgggtgtgcacacacatct---ggc	378
Db	485	gcttcgaagtacacatgcacgcgtgcctggacggggcggtgggtgtgatcccttgatgcagca	544
QY	379	tcgcagatgcgcggccgcgcgtctcgtgtgcggcccaacccgcggcgcgctgtagcatacttggc	438
Db	545	tggagcttccgtctccacagccctgtacgttcccttccccgagagaggtcaaacctccggga	604
QY	439	actgcctgtgagcagtggtggtatgtgaggaacgaacgcgaagaggccacgcgaagaccgcaccc	498
Db	605	aattcctcgaagagatgggtgtgtgacacgcccacgaagaccaaaccgttggtggcctgcc	664
QY	499	gtgcacacagagagccttcgtagtctgtgtgtagtggatggagatgacaaagga-----act	552
Db	665	tcggggctttaccgactggaagacacgctttggcccccagaccacctaattgattgaagccaaatt	724
QY	553	gcatagcctacacagagcccccgtggagccctgtctccacacagctgcgcgtgggtgtccca	612
Db	725	gccctgtgtccaaacacacagagtgagagccgccttctccaaagacctgtgggatgtggcatctcca	784
QY	613	ctcggatctccaatgttaaagcccccagtgctgtgccttgacgaagagagcgcgcctctgcacat	672
Db	785	ccccgggttaccgaatgacaaacccctcctgcagagctgaagaaagacagagccgcctgagcatgg	844
QY	673	tggcgccatgtgagtgtgagactccatacacatccatgaagcagaggaagaagtgtcgtctg	722
Db	845	tcagagccttgggaagcgtacctggaaagaaacattttaaaggaaggaagaaagtgcatccgta	904
QY	733	tgtaaccagccagagcagatcatalcatalcaacttcacactctggcggtgcgtacagcaacagctcct	792
Db	905	ctcccaaaaatttccaaagcgttatcaagtttagagctttcttgctgtcacacagcatgaagcat	964
QY	793	atcaacccaaagctatgtgaggttctgcatalygaacaaataggtctgtgcattccctacaaagtcta	852
Db	965	accgagctaaatttctgmgagatattgtaaccgacggccgactgtgcacccccacagaaacca	1024
QY	853	agaactatcgacgtgtcctctccagtgctccctgtagtggcttggtgcttcctccgcaggtcctat	912
Db	1025	ccacccttcggcgtgaggttcaagtgcccttgacggcgagctatcamaagaaacatgtatgt	108
QY	913	ggattaatgcctgtctctcgtataccttgagctgtgtagaaatcccaatgagatctttgtact	972
Db	1085	tcatacaaacctgttgcttgccttcatttaccatctgtcccgagacaaatgacatctttgaattgc	114
QY	973	tggaatccctacccctgagctcttcagaatattgc	1003
Db	1145	tgtactacagaaagatgtacggagacatggc	1175
RESULT	9		
LOCUS	132210	2075 bp	DNA
DEFINITION	Sequence 1 from patent US 5585270.		PAT
ACCESSION	132210		
VERSION	132210.1	GI:1823001	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 2075)		
AUTHORS	Giendordorf, G.R. and Bradham, D.M. Jr.		

TITLE	Polynucleotides encoding connective tissue growth factor
JOURNAL	Patent: US 5585370-A 1 17-DEC-1996;
FEATURES	Location/Qualifiers
SOURCE	1..2075
	/organism="unknown"
BASE COUNT	491 a 558 c 546 g 480 t
ORIGIN	.

[illegible]

```

Oy 973 tgggaatccctaccctgactctcagaattgc 1003
Db 1145 TGTACTACAGAGAATGTACGAGACATGGC 1175

RESULT 10
HOMCONGRO LOCUS HUMCONGRO 2075 bp mRNA PRI 05-MAR-1996
DEFINITION Human connective tissue growth factor, complete cds.
ACCESSION M92934 M36965 S56201
VERSION M92934.1 GI:180923
KEYWORDS growth factor; mitogen.
SOURCE Homo sapiens (Library: lambda gt11) connective cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2075)
AUTHORS Bradham,D.M., Igarashi,A., Potter,R.L. and Grotendorst,G.R.
TITLE Connective tissue growth factor: a cysteine-rich mitogen secreted
by human vascular endothelial cells is related to the SRC-induced
immediate early gene product CEF-10
JOURNAL J. Cell Biol. 114 (6), 1285-1294 (1991)
MEDLINE 91373462
REFERENCE 2 (bases 1 to 2075)
AUTHORS Grotendorst,G.R.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-1990) Gary R. Grotendorst, Department of Cell
Biology and Anatomy, University of Miami School of Medicine, Miami,
FL, 33136, USA

FEATURES
Source location/Qualifiers
1..2075
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="endothelial"
/tissue_type="connective"
/tissue_lib="lambda gt11"
1..2075
130..192
130..1179
/codon_start=1
/product="connective tissue growth factor"
/protein_id="AAA91279.1"
/db_xref="GI:180924"
/translating="MTASMEFVRYAVVLLACSRPVGONCSGPCRCPDEPARCP
AGVSLVDGCGCCFQCANOLGELCTERDPCDPRHKLFDGSPARKIGVCIANKGAP
CIFGCTVYRSGESFQSCYKQCTCDGAVGCPICSMYVRLSPCPPRKTLKPGK
CEWVCDEPKDQTVVGPALAAAYRLDPTGPDPTMIRANLIVOTITMSKSGKSGIS
TRVTNDNASCRLEKOSRLCMVRPEADILENKKKIKIKIRPKIKIKELSGCTSM
KTYRAKFGVCTDGGCTCPHRTTTPVEFKCPDGEVKNKMMFINTCACHYKCPEDND
IFESLYIRKMTGDMA"
193..1176
mat_peptide /product="connective tissue growth factor"
BASE COUNT 491 a 558 c 546 g 480 t
ORIGIN

Query Match 6.7%; Score 243.4; DB 85; Length 2075;
Best Local Similarity 56.2%; Pred. No. 2.4e-47;
Matches 523; Conservative 0; Mismatches 396; Indels 12; Gaps 3;

```

```

Db 425 ATGGCTCCCTCCGACATCTTCGGTGTACGGTATACCGACGAGATCTTCCAGAGCA 484
Oy 322 actgcaagtaaacacgctgcatcgacgagcgcggtggtgcaacacactgt---gcc 378
Db 485 GCTGCAGTACAGATGACAGTCTGACGAGGCGGCGGTGGCTGCAATGCCCTGTGACAGA 544
Oy 379 tcgagtgcgcccccgagctctctgtgccccccaccccgcgcggtgagatatactgccc 438
Db 545 TGGAGTTCGTGTGCCCGACGCTGTGACCTGCTCCCGAGGAGGTCAAGCTGCCGGGA 604
Oy 439 actgctgtgagcagtggtggtatgtgagagcagcagcagagcgagcagagacacccc 498
Db 605 AATGCTGCAGGAGTGGGGTGTGACGAGCCCAAGACCAACCGTGTGGCTGCC 664
Oy 499 gtgacacagagccttcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 552
Db 665 TGGCGGCTTACCGAGTGGAGACAGACGTTTGGCCGACACCACTATGATTAAGCCACT 724
Oy 553 gataagctacacagagcccttgagagcccttgctccacacagctgagcggtggtgtcca 612
Db 725 GCGTGTCAGACACAGAGAGTGGAGCGGCTGTCCAGACCTGTGGGATGGGCATCTCCA 784
Oy 613 ctgagatctcgaatttaagcccgagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 672
Db 785 CCGGGTTACCAATGACAAAGCGCTCTCTGAGGCTAGAGAGAGAGCGCTGTGCATGG 844
Oy 673 tggcgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 732
Db 845 TCAGGCTTGCAGACCTGACCTGCGAGAGACATTAAGAGAGGCAAAAAGTCAATCCGTA 904
Oy 733 tgtacacgacagagagcatcactgaacttcaacttgcggtgtgtgtgtgtgtgtgtgt 792
Db 905 CTCGCCAAATCTCCAAAGCCTATCAAGTTTGTGCTGTGACACAGACATGAAGACAT 964
Oy 793 atcaacccagagacgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 852
Db 965 ACCGAGCTAAATCTGTGAGATGTATACGAGCGCGAGTGTGCAACCCCAAGAACCA 1024
Oy 853 agactatcgagctgtctccctcagctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 912
Db 1025 CCACCTGCGCGGTGAGTTCAGATTCAGTCCCTGACGCGGAGGTCAATGAAGAACATGAT 1084
Oy 913 gattaatgctgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 972
Db 1085 TATCAAGACCTGCTGCTCCATTAACACTGTCCGAGACATATGACATTTGAATCGC 1144
Oy 973 tggatctcactcctgactctcagaattgc 1003
Db 1145 TGTACTACAGAGAATGTACGAGACATGGC 1175

RESULT 11
HSCGF LOCUS HSCGF 2312 bp mRNA PRI 10-MAR-1997
DEFINITION H.sapiens mRNA for connective tissue growth factor.
ACCESSION X78947
VERSION X78947.1 GI:474933
KEYWORDS connective tissue; growth factor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2312)
AUTHORS Oemar,B.S., Werner,A., Garrier,J.M., Do,D.D., Godoy,N., Nauck,M.,
Matz,W., Rupp,J., Pech,M. and Luecher,T.F.
TITLE Human connective tissue growth factor is expressed in advanced
atherosclerotic lesions
JOURNAL Circulation 95 (4), 831-839 (1997)
MEDLINE 97207446
REFERENCE 2 (bases 1 to 2312)
AUTHORS Oemar,B.S.
TITLE Direct Submission

```





Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L3009  
Center clone name: 9\_F\_9

\* NOTE: This record contains 152 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
799 898: contig of 798 bp in length  
899 1659: contig of 771 bp in length  
1670 1769: gap of 100 bp  
1770 2540: contig of 771 bp in length  
2541 2640: gap of 100 bp  
2641 3386: contig of 756 bp in length  
3397 3496: gap of 100 bp  
3497 4286: contig of 790 bp in length  
4287 4386: gap of 100 bp  
4387 5167: contig of 781 bp in length  
5168 5267: gap of 100 bp  
5268 6055: contig of 788 bp in length  
6056 6155: gap of 100 bp  
6156 6940: contig of 785 bp in length  
6941 7040: gap of 100 bp  
7041 7834: contig of 794 bp in length  
7835 7934: gap of 100 bp  
7935 8740: contig of 806 bp in length  
8741 8840: gap of 100 bp  
8841 9619: contig of 779 bp in length  
9620 9719: gap of 100 bp  
9720 10498: contig of 779 bp in length  
10499 10598: gap of 100 bp  
10599 11352: contig of 754 bp in length  
11353 11452: gap of 100 bp  
11453 12245: contig of 793 bp in length  
12246 12345: gap of 100 bp  
12346 13127: contig of 782 bp in length  
13128 13227: gap of 100 bp  
13228 14019: contig of 792 bp in length  
14020 14119: gap of 100 bp  
14120 14915: contig of 796 bp in length  
14916 15015: gap of 100 bp  
15016 15784: contig of 769 bp in length  
15785 15884: gap of 100 bp  
15885 16669: contig of 785 bp in length  
16670 16769: gap of 100 bp  
16770 17539: contig of 770 bp in length  
17540 17639: gap of 100 bp  
17640 18432: contig of 793 bp in length  
18433 18532: gap of 100 bp  
18533 19296: contig of 764 bp in length  
19297 19396: gap of 100 bp  
19397 20173: contig of 777 bp in length  
20174 20273: gap of 100 bp  
20274 21046: contig of 773 bp in length  
21047 21146: gap of 100 bp  
21147 21933: contig of 787 bp in length  
21934 22033: gap of 100 bp  
22034 22823: contig of 790 bp in length  
22824 22923: gap of 100 bp  
22924 23706: contig of 783 bp in length  
23707 23806: gap of 100 bp  
23807 24579: contig of 773 bp in length  
24580 24679: gap of 100 bp

24690 25481: contig of 802 bp in length  
25482 25581: gap of 100 bp  
25582 26374: contig of 793 bp in length  
26375 26474: gap of 100 bp  
26475 27247: contig of 773 bp in length  
27248 27347: gap of 100 bp  
27348 28133: contig of 786 bp in length  
28134 28233: gap of 100 bp  
28234 29016: contig of 783 bp in length  
29017 29116: gap of 100 bp  
29117 29882: contig of 766 bp in length  
29883 29982: gap of 100 bp  
29983 30718: contig of 736 bp in length  
30719 30818: gap of 100 bp  
30819 31594: contig of 776 bp in length  
31595 31694: gap of 100 bp  
31695 32454: contig of 760 bp in length  
32455 32554: gap of 100 bp  
32555 33312: contig of 758 bp in length  
33313 33412: gap of 100 bp  
33413 34176: contig of 764 bp in length  
34177 34276: gap of 100 bp  
34277 35068: contig of 792 bp in length  
35069 35168: gap of 100 bp  
35169 35960: contig of 792 bp in length  
35961 36060: gap of 100 bp  
36061 36833: contig of 773 bp in length  
36834 36933: gap of 100 bp  
36934 37703: contig of 770 bp in length  
37704 37803: gap of 100 bp  
37804 38592: contig of 789 bp in length  
38593 38692: gap of 100 bp  
38693 39480: contig of 788 bp in length  
39481 39580: gap of 100 bp  
39581 40371: contig of 791 bp in length  
40372 40471: gap of 100 bp  
40472 41251: contig of 780 bp in length  
41252 41351: gap of 100 bp  
41352 42138: contig of 787 bp in length  
42139 42238: gap of 100 bp  
42239 43010: contig of 772 bp in length  
43011 43110: gap of 100 bp  
43111 43884: contig of 774 bp in length  
43885 43984: gap of 100 bp  
43985 44768: contig of 784 bp in length  
44769 44868: gap of 100 bp  
44869 45643: contig of 775 bp in length  
45644 45743: gap of 100 bp  
45744 46510: contig of 767 bp in length  
46511 46610: gap of 100 bp  
46611 47388: contig of 778 bp in length  
47389 47488: gap of 100 bp  
47489 48270: contig of 782 bp in length  
48271 48370: gap of 100 bp  
48371 49160: contig of 790 bp in length  
49161 49260: gap of 100 bp  
49261 50035: contig of 775 bp in length  
50036 50135: gap of 100 bp  
50136 50916: contig of 781 bp in length  
50917 51016: gap of 100 bp  
51017 51801: contig of 785 bp in length  
51802 51901: gap of 100 bp  
51902 52680: contig of 789 bp in length  
52681 52790: gap of 100 bp  
52791 53578: contig of 788 bp in length  
53579 53678: gap of 100 bp  
53679 54447: contig of 769 bp in length  
54448 54547: gap of 100 bp  
54548 55343: contig of 796 bp in length  
55344 55443: gap of 100 bp  
55444 56213: contig of 770 bp in length  
56214 56313: gap of 100 bp  
56314 57080: contig of 767 bp in length











GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2001, 10:15:19 ; Search time 152.1 Seconds  
(without alignments)  
9034.675 Million cell updates/sec

Title: US-09-325-019-1  
Perfect score: 3658  
Sequence: 1 cgcacttaccacccagctcca.....aaaaaaaaaaaaaaaa 3658

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

N.Geneseq.36.\*  
1: /SID56/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SID56/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SID56/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SID56/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SID56/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SID56/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SID56/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SID56/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SID56/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SID56/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SID56/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SID56/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SID56/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SID56/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SID56/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SID56/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SID56/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SID56/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SID56/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SID56/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3658	100.0	3658	21	Human connective t
2	2611.8	71.4	2830	20	Human WISP-1 prote
3	1193	32.6	1403	20	Human WISP-1 clone
4	1004.4	27.5	1101	20	Human WISP-1 prote
5	815.4	22.3	1766	20	Mouse WISP-1 prote
6	773.8	21.2	1202	20	Human WISP-1 clone
7	747.6	20.4	1140	20	Mouse WISP-1 prote
8	735	20.1	1183	20	Human WISP-1 clone
9	459.8	12.6	1101	20	Human WISP-1 clone
10	404	11.0	546	20	Human WISP-1 clone
11	403.4	11.0	683	20	Human WISP-1 clone
12	355.8	9.7	540	21	Human CTGF-4 relat

c	13	316.4	8.6	539	21	258615	Human CTGF-4 relat
	14	295.4	8.1	311	21	258616	Human CTGF-4 relat
	15	277.8	7.6	1062	20	211246	Human connective t
	16	255.2	7.0	693	20	276491	Human WISP-1 clone
	17	243.4	6.7	2075	16	T04226	Connective tissue
	18	243.4	6.7	2075	16	T45360	Human connective t
	19	243.4	6.7	2075	18	T51234	Connective tissue
	20	243.4	6.7	2075	19	V38085	Human connective t
	21	243.4	6.7	2075	20	X61317	Human connective t
	22	225.4	6.2	2338	20	X90030	Rat connective tis
	23	220.4	6.0	2330	13	Q26422	Gene for beta-1G-M
	24	218.8	6.0	2267	18	T94700	Murine Flsp12 CDNA
	25	199.6	5.5	1335	20	X76497	Human WISP-3 prote
c	26	199.6	5.5	1335	20	X76498	Human WISP-3 prote
	27	198	5.4	1142	20	211245	Human connective t
	28	198	5.4	1183	21	A06928	Human growth facto
	29	198	5.4	1212	20	X76499	Human WISP-3 prote
c	30	183.4	5.0	1212	20	X76500	Human WISP-3 prote
	31	183.4	5.0	197	21	258617	Human CTGF-4 relat
	32	167	4.6	1975	14	Q36031	Chicken nov coding
	33	154.6	4.2	484	21	258618	Human CTGF-4 relat
	34	130.4	3.6	684	14	Q36032	Chicken nov gene f
	35	129.6	3.5	738	20	X76501	Human WISP-2 prote
c	36	129.6	3.5	750	20	X76487	Human WISP-2 prote
	37	129.6	3.5	1257	20	X28435	EGF-like homologue
	38	129.6	3.5	1265	21	A30048	Human PRO261 nucle
	39	129.6	3.5	1285	19	V29260	Human connective t
	40	129.6	3.5	1293	20	X76486	Human WISP-2 prote
	41	129.6	3.5	1522	20	X16595	Human growth facto
	42	124.4	3.4	2028	13	Q26421	Gene for beta-1G-M
	43	123.6	3.4	841	20	X76502	Human WISP-2 prote
	44	120.8	3.3	531	21	A06931	Human growth facto
	45	111.6	3.1	236	21	258619	Human CTGF-4 relat

#### ALIGNMENTS

RESULT	1	
ID	258613	standard; CDNA: 3658 BP.
AC	258613;	
XX		
DT	11-APR-2000	(first entry)
DE	Human connective tissue growth factor-4 (CTGF-4) encoding CDNA.	
XX		
KW	Connective tissue growth factor-4; CTGF-4; human; immune system disorder;	
KW	hematopoietic disorder; autoimmune disorder; diabetes mellitus; asthma;	
KW	respiratory disorder; inflammation; hyperproliferative disorder;	
KW	infection; central nervous system disease; Alzheimer's disease; AIDS;	
KW	food additive; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	3..4010
FT		/*tag- a
FT		/Product- "CTGF-4"
PN	W09962927-AL	
XX		
PD	09-DEC-1999.	
XX		
PF	03-JUN-1999;	99MO-US12150.
XX		
PR	05-JUN-1998;	98US-0088320.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Young PE;	
XX		

DR WPI, 2000-147042/13.  
DR P-SDS; Y59247.  
XX  
PT New isolated connective tissue growth factor-4, used for treating e.g.  
PT cancers -  
XX  
PS Claim 4; Fig 1A-E; 196pp; English.

XX The invention provides an isolated human connective tissue growth factor  
CC -4 (CTGF-4) polypeptide. The CTGF-4 cDNA is deposited under ATCC No.  
CC 209816. The CTGF-4 protein can be expressed by standard recombinant  
CC methodology. The polypeptides can be used for preventing, treating or  
CC ameliorating a medical condition. They may be useful in treating  
CC deficiencies or disorders of the immune system, by activating or  
CC inhibiting the proliferation, differentiation, or mobilization  
CC (chemotaxis) of immune cells, treating or detecting deficiencies or  
CC disorders of hematopoietic cells (e.g. blood protein disorders, ataxia  
CC telangiectasia, HIV infection, Di George syndrome, anemia or  
CC hemoglobinuria), to modulate hemostatic (the stopping of bleeding) or  
CC thrombolytic activity (clot formation) (e.g. blood coagulation disorders,  
CC blood platelet disorders, or wounds resulting from trauma, or surgery),  
CC in treating or detecting autoimmune disorders (e.g. Addison's disease,  
CC rheumatoid arthritis, allergic encephalomyelitis, Goodpasture syndrome,  
CC multiple sclerosis, purpura, Reiter's disease, Guillain-Barre syndrome,  
CC systemic lupus erythematosus, insulin dependent diabetes mellitus or  
CC autoimmune inflammatory eye disease), treating asthma (particularly  
CC allergic asthma) or other respiratory problems (e.g. anaphylaxis,  
CC hypersensitivity to an antigenic molecule or blood group  
CC incompatibility), to treat and/or prevent organ rejection or graft-versus  
CC host disease (GVHD), to modulate inflammation (septic shock, sepsis,  
CC arthritis, nephritis, cytokine or chemokine induced lung injury,  
CC inflammatory bowel disease, Crohn's disease, or resulting from over  
CC production of cytokines), to treat hyperproliferative disorders,  
CC including neoplasms in the abdomen, bone, breast, digestive system,  
CC liver, pancreas, peritoneum, endocrine glands, eye, head and neck, soft  
CC nervous (central and peripheral), lymphatic system, pelvic, skin, soft  
CC tissue, spleen, thoracic and urogenital, hypergammaglobulinemia,  
CC lymphoproliferative disorders, Waldenstrom's macroglobulinemia,  
CC sarcoidosis), to treat or detect infectious agents, e.g. viruses (e.g.  
CC arthritis, bronchiolitis, encephalitis, eye infections, chronic fatigue  
CC syndrome, hepatitis, meningitis, AIDS, pneumonia, chickenpox, measles,  
CC mumps, parainfluenza, rabies, the common cold, polio, leukemia, rubella,  
CC sexually transmitted diseases, or skin diseases) bacterial or fungal  
CC agents (e.g. bacteremia, endocarditis, eye infections, gingivitis,  
CC opportunistic infections, respiratory tract infections, Lyme disease,  
CC cat-scratch disease, paratyphoid fever, food poisoning, pneumonia,  
CC gonorrhea and sexually transmitted diseases, meningitis, tuberculosis,  
CC lupus, gangrene, tetanus, rheumatic fever, urinary tract infections,  
CC wound infections), parasitic agents (e.g. scabies, dysentery, liver  
CC disease, malaria, toxoplasmosis), to differentiate, proliferate and  
CC attract cells, leading to the regeneration of tissues (e.g. repair,  
CC replace or protect tissue in wounds, burns, incisions or ulcers,  
CC osteoporosis, osteoarthritis, periodontal disease, liver failure,  
CC surgery, cosmetic plastic surgery, reperfusion injury) to proliferate and  
CC differentiate nerve cells (e.g. spinal cord disorders, head trauma,  
CC cerebrovascular disease and stroke). Localized neuropathies and central  
CC nervous system diseases (e.g. Alzheimer's disease, Parkinson's disease,  
CC Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager  
CC syndrome). They may also increase or decrease the differentiation or  
CC proliferation of embryonic stem cells and hematopoietic lineage, may be  
CC used to modulate mammalian characteristics such as body height, weight,  
CC hair color, eye color, skin, percentage of adipose tissue, pigmentation,  
CC size, and shape, to modulate mammalian metabolism affecting catabolism,  
CC anabolism, processing, utilization and storage of energy, to change a  
CC mammal's mental state or physical state by influencing biohythms,  
CC cardiac rhythms, circadian rhythms, depression (including depressive  
CC disorders), tendency for violence, tolerance for pain, libido, memory,  
CC capabilities, hormonal or endocrine levels, appetite, libido, memory,  
CC stress, or other cognitive qualities, as a food additive or preservative,  
CC such as to increase or decrease storage capabilities, fat content, lipid,  
CC protein, carbohydrate, vitamins, minerals, cofactors or other nutritional  
CC components. Mutations in the PNs or the presence or amount of expression  
CC or activity of the polypeptides can be used for diagnosing a pathological

CC condition or a susceptibility to a pathological condition. The CTGF-4  
CC polypeptides can also be used for identifying binding partners. The  
CC products can also be used for producing transgenic animals. The present  
CC sequence represents the CTGF-4 encoding cDNA.  
XX  
SQ Sequence 3658 BP; 984 A; 851 C; 841 G; 982 T; 0 other;

Query Match 100.0%; Score 3658; DB 21; Length 3658;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 cggacttaccacagctccacagctgagagacacccctccacagcccaatctcgaagtgc 60  
DB 1 cggacttaccacagctccacagctgagagacacccctccacagcccaatctcgaagtgc 60  
OY 61 catgtatgagccgcatccaccccgctgcccgtgggggtcagccatcagatg 120  
DB 61 catgtatgagccgcatccaccccgctgcccgtgggggtcagccatcagatg 120  
OY 121 gctgtatgctgtatgagctgctcagcagagcttgaggacactgacagagctgcga 180  
DB 121 gctgtatgctgtatgagctgctcagcagagcttgaggacactgacagagctgcga 180  
OY 181 tctgtaccacccacgggggctctactgtgtactacagcggggagccgaggtacgaa 240  
DB 181 tctgtaccacccacgggggctctactgtgtactacagcggggagccgaggtacgaa 240  
OY 241 taagatgt 300  
DB 241 taagatgt 300  
OY 301 acggcagctctccacagctcactacagctacagctacagctacagctacagctacag 360  
DB 301 acggcagctctccacagctcactacagctacagctacagctacagctacagctacag 360  
OY 361 gctgtac 420  
DB 361 gctgtac 420  
OY 421 gctgtac 480  
DB 421 gctgtac 480  
OY 481 cagcagac 540  
DB 481 cagcagac 540  
OY 541 ggcacaggaac 600  
DB 541 ggcacaggaac 600  
OY 601 tgggt 660  
DB 601 tgggt 660  
OY 661 ggccttcgaacttcgagcagctgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720  
DB 661 ggccttcgaacttcgagcagctgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 720  
OY 721 agtgcctgt 780  
DB 721 agtgcctgt 780  
OY 781 ggcacagctctctac 840  
DB 781 ggcacagctctctac 840  
OY 841 cctacagctacagctacagctacagctacagctacagctacagctacagctacagctac 900  
DB 841 cctacagctacagctacagctacagctacagctacagctacagctacagctacagctac 900  
OY 901 gccaggtcctatgtatgtatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960

Db 901 gccagttccatgataatgctgtctctgtaacctgagctgtagaatcccaatgaca 960  
QY 961 tctttctgacttggaatcttaacctgacttccagaatctgccaactgagcgacaa 1020  
Db 961 tctttctgacttggaatcttaacctgacttccagaatctgccaactgagcgacaa 1020  
QY 1021 attctggtcttgaggactaaaccaatgctgtggaagcagtcagcccttctgccaataa 1080  
Db 1021 attctggtcttgaggactaaaccaatgctgtggaagcagtcagcccttctgccaataa 1080  
QY 1081 ctttccaccaatgagccttagttacccctgtaacctgagcccttgccctccattctgtctc 1140  
Db 1081 ctttccaccaatgagccttagttacccctgtaacctgagcccttgccctccattctgtctc 1140  
QY 1141 aaccatctcaaatgagcctgtaagtgtgtctgctcagcgcccaatgctatgattctctctga 1200  
Db 1141 aaccatctcaaatgagcctgtaagtgtgtctgctcagcgcccaatgctatgattctctctga 1200  
QY 1201 tatcatccagcatctactctaaagaaaatgctgtctctagctgttctctgactacacc 1260  
Db 1201 tatcatccagcatctactctaaagaaaatgctgtctctagctgttctctgactacacc 1260  
QY 1261 aagcctgataccagccttcccaagtcactagaagtcctgtctgacttctgcttaattcccaa 1320  
Db 1261 aagcctgataccagccttcccaagtcactagaagtcctgtctgacttctgcttaattcccaa 1320  
QY 1321 gaataggaatcagtgactttaaatacctaattctctctctagatgccaacacaa 1380  
Db 1321 gaataggaatcagtgactttaaatacctaattctctctctagatgccaacacaa 1380  
QY 1381 gaactcttggtccattcagaatgaatggaatttggaacaaatggaataacttaatt 1440  
Db 1381 gaactcttggtccattcagaatgaatggaatttggaacaaatggaataacttaatt 1440  
QY 1441 ttggaagcctgccaagagtgactgtaagttaattctgtaagtcagcgacacaaactatc 1500  
Db 1441 ttggaagcctgccaagagtgactgtaagttaattctgtaagtcagcgacacaaactatc 1500  
QY 1501 ctgataccaaatgataatgacactcaagtcacaaacatttgcgaagtgaattgaatag 1560  
Db 1501 ctgataccaaatgataatgacactcaagtcacaaacatttgcgaagtgaattgaatag 1560  
QY 1561 ttgcttaatttgatttttaatgaaagtgtatccaattactgagcattgtttgagtt 1620  
Db 1561 ttgcttaatttgatttttaatgaaagtgtatccaattactgagcattgtttgagtt 1620  
QY 1621 aagttctctccacccctacacgtgtaaggtgacagattagtttgcacagcagaat 1680  
Db 1621 aagttctctccacccctacacgtgtaaggtgacagattagtttgcacagcagaat 1680  
QY 1681 aaaaattgataaacaatctctgtgtgtaaggaagcccccgatgaatactccagaagacag 1740  
Db 1681 aaaaattgataaacaatctctgtgtgtaaggaagcccccgatgaatactccagaagacag 1740  
QY 1741 gaaagtccagccattcagaagaaagcaatgacttccacacagaaacagctgtgactg 1800  
Db 1741 gaaagtccagccattcagaagaaagcaatgacttccacacagaaacagctgtgactg 1800  
QY 1801 gcaagagccttgaggcaggttgagcagcctctccctgtaactctccctgtccctgctggg 1860  
Db 1801 gcaagagccttgaggcaggttgagcagcctctccctgtaactctccctgtccctgctggg 1860  
QY 1861 gtccataggaatgtgtaagcctctgagactgagcctgtctgagccctgagagtgagcct 1920  
Db 1861 gtccataggaatgtgtaagcctctgagactgagcctgtctgagccctgagagtgagcct 1920  
QY 1921 ggaacactcctctactcttcaagagccttgaagagcccaagctgcaagccctgcaagcc 1980  
Db 1921 ggaacactcctctactcttcaagagccttgaagagcccaagctgcaagccctgcaagcc 1980  
QY 1981 acgaaatgacaaagacagttcaggtagagggtgtgtgtaaaccaagaggtgtgccc 2040  
Db 1981 acgaaatgacaaagacagttcaggtagagggtgtgtgtaaaccaagaggtgtgccc 2040  
Db 1981 actgaatgacaaagacagttcaggttaggggtgtgtgtgtaaaccaagaggtgtgccc 2040  
QY 2041 ttgtagcagcctggggtgaacctctgagctgagagctgttggaacccagggggcccggt 2100  
Db 2041 ttgtagcagcctggggtgaacctctgagctgagagctgttggaacccagggggcccggt 2100  
QY 2101 gtccagacacacatctctcagaagactatccacagccttctgtctctgagcacaatg 2160  
Db 2101 gtccagacacacatctctcagaagactatccacagccttctgtctctgagcacaatg 2160  
QY 2161 gccagtttcttgtaggaagatggaagttbaacagttgtttagaaaacagaatagactta 2220  
Db 2161 gccagtttcttgtaggaagatggaagttbaacagttgtttagaaaacagaatagactta 2220  
QY 2221 ataaagtttaagcctggaagaggttgaagctaaagaaaaggtttgtttagtaataat 2280  
Db 2221 ataaagtttaagcctggaagaggttgaagctaaagaaaaggtttgtttagtaataat 2280  
QY 2281 cagcctatctatctgataatgaagaaataatactactgtttagaatctcttctta 2340  
Db 2281 cagcctatctatctgataatgaagaaataatactactgtttagaatctcttctta 2340  
QY 2341 gggccttctctgtgccagacatgtctcagtgcttgcagtgatagctcactgaact 2400  
Db 2341 gggccttctctgtgccagacatgtctcagtgcttgcagtgatagctcactgaact 2400  
QY 2401 tcaagcaatglttgaaagttcccatattattctgtcttccaaatgtaacggag 2460  
Db 2401 tcaagcaatglttgaaagttcccatattattctgtcttccaaatgtaacggag 2460  
QY 2461 ctcatagaggtgagaaacccaacccagagtcacccaggttgtagctgggaagttgagat 2520  
Db 2461 ctcatagaggtgagaaacccaacccagagtcacccaggttgtagctgggaagttgagat 2520  
QY 2521 tcaagatcgaatattggaactgtctttaaaccacatatttcccccgtttttagagctcca 2580  
Db 2521 tcaagatcgaatattggaactgtctttaaaccacatatttcccccgtttttagagctcca 2580  
QY 2581 aatgtctcagaataaggaataacatctgcaataaagcttgaatttttaatgtaatttcc 2640  
Db 2581 aatgtctcagaataaggaataacatctgcaataaagcttgaatttttaatgtaatttcc 2640  
QY 2641 ctcttaagcttctctagcctcttccaaaagcgggaatctgtaatttctgataatt 2700  
Db 2641 ctcttaagcttctctagcctcttccaaaagcgggaatctgtaatttctgataatt 2700  
QY 2701 aggtgcttaagcattccaaaatacagtggaacacaaatccaggaatcccccgttagct 2760  
Db 2701 aggtgcttaagcattccaaaatacagtggaacacaaatccaggaatcccccgttagct 2760  
QY 2761 attccctcttcccatcggaacacagctctcaatcaacatttaagaatgattctgtttac 2820  
Db 2761 attccctcttcccatcggaacacagctctcaatcaacatttaagaatgattctgtttac 2820  
QY 2821 ccaatgctgataatgaaagtgtgtgagttatccaaaggaatctctgtaaggtgtcaga 2880  
Db 2821 ccaatgctgataatgaaagtgtgtgagttatccaaaggaatctctgtaaggtgtcaga 2880  
QY 2881 gaattctcctaaaggggaagagcgtggaataacatccctctactgtgacctcccaaa 2940  
Db 2881 gaattctcctaaaggggaagagcgtggaataacatccctctactgtgacctcccaaa 2940  
QY 2941 cctagtcaggtggaaggtatagagtggtgtctcaatgaatcttgatgaagaagagct 3000  
Db 2941 cctagtcaggtggaaggtatagagtggtgtctcaatgaatcttgatgaagaagagct 3000  
QY 3001 gtgcaggtgtccctactttatctcgaagcctcttcttccaaagctacagaaatagaa 3060  
Db 3001 gtgcaggtgtccctactttatctcgaagcctcttcttccaaagctacagaaatagaa 3060  
QY 3061 ttcttaacagtcacaaatttatataggtgtcttagcaaaagagacattatattccaa 3120  
Db 3061 ttcttaacagtcacaaatttatataggtgtcttagcaaaagagacattatattccaa 3120  
QY 3061 ttcttaacagtcacaaatttatataggtgtcttagcaaaagagacattatattccaa 3120  
Db 3061 ttcttaacagtcacaaatttatataggtgtcttagcaaaagagacattatattccaa 3120

QY 3121 gaattgcatatgatacatatatttgatagcgtatgcaagaattctgtatanaa 3180  
 |||||||  
 Db 3121 gaattgcatatgatacatatatttgatagcgtatgcaagaattctgtatanaa 3180  
 QY 3181 gagaattcactcatatgataatctctctctgtaagtcagttgtaattctgtatcttc 3240  
 |||||||  
 Db 3181 gagaattcactcatatgataatctctctctgtaagtcagttgtaattctgtatcttc 3240  
 QY 3241 gagaattcactcatatgataatctctctctgtaagtcagttgtaattctgtatcttc 3300  
 |||||||  
 Db 3241 gagaattcactcatatgataatctctctctgtaagtcagttgtaattctgtatcttc 3300  
 QY 3301 aaattcagtggaaccccaagagcaggttcttactgtaatttgatctcaatttggaac 3360  
 |||||||  
 Db 3301 aaattcagtggaaccccaagagcaggttcttactgtaatttgatctcaatttggaac 3360  
 QY 3361 tgaacttaagcttcaaaagatcatatgaagtcagttgtaagtaactagattatctgc 3420  
 |||||||  
 Db 3361 tgaacttaagcttcaaaagatcatatgaagtcagttgtaagtaactagattatctgc 3420  
 QY 3421 aattgagtggaagagtggtgctcaagataaccagtcagtcagttggttggaacaggaact 3480  
 |||||||  
 Db 3421 aattgagtggaagagtggtgctcaagataaccagtcagtcagttggttggaacaggaact 3480  
 QY 3481 attctatgcagagcccgagagcgcaagtcacaaaccatgggttggaagccatggaagga 3540  
 |||||||  
 Db 3481 attctatgcagagcccgagagcgcaagtcacaaaccatgggttggaagccatggaagga 3540  
 QY 3541 aattgagtggaagagtggtgctcaagataaccagtcagtcagttggttggaacaggaact 3600  
 |||||||  
 Db 3541 aattgagtggaagagtggtgctcaagataaccagtcagtcagttggttggaacaggaact 3600  
 QY 3601 ctgattcaataaataatcaagataataataatggaagaaataaataaataaataaataa 3658  
 |||||||  
 Db 3601 ctgattcaataaataatcaagataataataatggaagaaataaataaataaataaataa 3658  
 RESULT 2  
 X76482  
 ID X76482 standard: DNA; 2830 BP.  
 AC X76482;  
 XX  
 DT 06-AUG-1999 (first entry)  
 XX  
 DE Human WISP-1 protein nucleotide sequence SEQ ID NO:1.  
 XX  
 KW WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
 KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
 KW leukemia; lymphoid malignancy; haematopoiesis-related disorder;  
 KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
 KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
 KW connective tissue disorder; catabolic state; inflammation;  
 KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO9921998-A1.  
 PD 06-MAY-1999.  
 XX  
 PF 29-OCT-1998; 98MO-US22991.  
 XX  
 PR 14-APR-1998; 98US-0081695.  
 PR 29-OCT-1997; 97US-0063704.  
 PR 03-FEB-1998; 98US-0073612.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
 XX Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
 XX

DR WPI: 1999-337420/28.  
 DR P-PSDB: Y17641.  
 PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
 XX  
 PS Example 3; Page 159-161; 284pp; English.  
 XX  
 CC The present invention describes Wnt-1 induced secreted polypeptides,  
 CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
 CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
 CC Products from the present invention can be used to treat WISP-related  
 CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
 CC products can be used to treat arteriosclerosis. The products can also be  
 CC used to treat other diseases e.g. benign and malignant tumours,  
 CC leukemia and lymphoid malignancies, neuronal, glial, astrocytal,  
 CC hypohalamic and other glandular, macropapal, epithelial, stromal, and  
 CC blastococell disorders, haematopoiesis-related disorders, tissue-growth  
 CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
 CC disorders, bone-related disorders such as osteoporosis, trauma such as  
 CC burns, incisions, and other wounds, connective tissue disorders,  
 CC catabolic states, testicular-related disorders, and inflammatory,  
 CC angiogenic and immunologic disorders including arteriosclerosis. The  
 CC products can also be used for detection and diagnosis especially of  
 CC individuals with neoplastic cell growth or proliferation. The products  
 CC can be used in the production of transgenic or knock-out animals.  
 CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
 CC cells.  
 XX  
 SO Sequence 2830 BP; 690 A; 736 C; 696 G; 708 T; 0 other;  
 Query Match 71.4%; Score 2611.8; DB 20; Length 2830;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2624; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
 QY 2 ggaattaccccccgcctcactggaagacacactctcaagcccccattctgcaagggcc 61  
 |||||||  
 Db 184 ggaattaccccccgcctcactggaagacacactctcaagcccccattctgcaagggcc 243  
 QY 62 atgtgagtgccgcgcacatcccccaccccgctgcgggtggttagctcatcaagattg 121  
 |||||||  
 Db 244 atgtgagtgccgcgcacatcccccaccccgctgcgggtggttagctcatcaagattg 303  
 QY 122 ctgtgagtgctgtaagatgtgcctcagcagcttggtggaacatgcaaggaagctggcat 181  
 |||||||  
 Db 304 ctgtgagtgctgtaagatgtgcctcagcagcttggtggaacatgcaaggaagctggcat 363  
 QY 182 ctgtgagtcacccacccgggctctcactgtgactacagcggggagccggggtacggaat 241  
 |||||||  
 Db 364 ctgtgagtcacccacccgggctctcactgtgactacagcggggagccggggtacggaat 423  
 QY 242 aggaagtggtgcaagagtggtgctggtggtgctgctcctgtagtgggtggtgcaacaac 301  
 |||||||  
 Db 424 aggaagtggtgcaagagtggtgctggtggtgctgctcctgtagtgggtggtgcaacaac 483  
 QY 302 cggccagctcttcacagcactacatgcaagtaacaactgacagctgacagcgcggtggg 361  
 |||||||  
 Db 484 cggccagctcttcacagcactacatgcaagtaacaactgacagctgacagcgcggtggg 543  
 QY 362 ctgacacacacatgtgctcctcagatgctgcccccggtctctgtgcccccaacccggcg 421  
 |||||||  
 Db 544 ctgacacacacatgtgctcctcagatgctgcccccggtctctgtgcccccaacccggcg 603  
 QY 422 cgtgagcatatcgtgcccactgctgtgagcagtggtgtagtgaagacgacgcaagagggc 481  
 |||||||  
 Db 604 cgtgagcatatcgtgcccactgctgtgagcagtggtgtagtgaagacgacgcaagagggc 663  
 QY 482 acgcaagacgcaaccccggtgacacagagagccttcagatgctgtggtgtagtgaagagcag 541  
 |||||||  
 Db 664 acgcaagacgcaaccccggtgacacagagagccttcagatgctgtggtgtagtgaagagcag 723  
 QY 542 gacaggaactgtagctacacaaagccctggagcccttgctcacaagctggcgct 601  
 |||||||







RESULT 4  
X76483/C  
ID X76483 standard; DNA: 1101 BP.  
XX  
XX X76483;  
DT 06-AUG-1999 (first entry)  
XX  
DE Human WISP-1 protein complementary nucleotide sequence SEQ ID NO:2.  
XX  
XX WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;  
KW connective tissue growth factor; cancer; melanoma; arteriosclerosis;  
KW leukemia; lymphoid malignancy; haematopoiesis-related disorder;  
KW tissue-growth disorder; skin disorder; desmoplasia; fibrotic lesion;  
KW kidney disorder; bone-related disorder; osteoporosis; trauma; burn;  
KW connective tissue disorder; catabolic state; inflammation;  
KW testicular-related disorder; angiogenesis; immunological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN MO9921998-A1.  
XX  
PD 06-MAY-1999.  
XX  
PF 29-OCT-1998; 98WO-US22991.  
XX  
PR 14-APR-1998; 98US-0081695.  
PR 29-OCT-1997; 97US-0063704.  
PR 03-FEB-1998; 98US-0073612.  
XX  
XX (GETH) GENENTECH INC.  
XX  
PI Botstein DA, Cohen RL, Goddard A, Gurney AL, Hillan K;  
PI Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;  
XX  
DR WPI; 1999-337420/28.  
XX  
PT New isolated Wnt-1 induced secreted polypeptides, WISP-1, 2 and 3  
PS  
XX Disclosure; Page 161-162; 284pp; English.  
XX  
XX The present invention describes Wnt-1 induced secreted polypeptides,  
CC WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2  
CC and WISP-3 have homology to connective tissue growth factor (CTGF).  
CC Products from the present invention can be used to treat WISP-related  
CC disorders such as breast, ovarian, and colon cancer or melanoma. The  
CC products can be used to treat arteriosclerosis. The products can also be  
CC used to treat other diseases e.g. benign and malignant tumours,  
CC leukemia and lymphoid malignancies, neuronal, glial, astrocytal,  
CC hypochromic and other glandular, macrophagal, epithelial, stromal, and  
CC blastocoele disorders, haematopoiesis-related disorders, tissue-growth  
CC disorders, skin disorders, desmoplasia, fibrotic lesions, kidney  
CC disorders, bone-related disorders such as osteoporosis, trauma such as  
CC burns, incisions, and other wounds, connective tissue disorders,  
CC catabolic states, testicular-related disorders, and inflammatory,  
CC angiogenic and immunologic disorders including arteriosclerosis. The  
CC products can also be used for detection and diagnosis especially of  
CC individuals with neoplastic cell growth or proliferation. The products  
CC can be used in the production of transgenic or knock-out animals.  
CC Antibodies can be used to induce death in WISP-1, 2 or 3 overexpressing  
CC cells.  
XX  
SQ Sequence 1101 BP; 219 A; 310 C; 354 G; 218 T; 0 other;

Query Match 27.5%; Score 1004.4; DB 20; Length 1101;  
Best Local Similarity 99.9%; Pred. NO. 5e-245;  
Matches 1005; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 ggaattacccagctcagctgagagacacccctcagcccccattctgcaagtgcc 61  
DB 1006 GGAATTACCCAGCTCAGCTGAGAGACACCTCTCAGCCCCCAATTCTGCAAGTGCC 947

QY 62 atgtgagtgccgcacatcccccacccgcctgcccctggggggtcagcctcaccagatg 121  
DB 946 ATGTGAGTGCCTCCGCATCCCCACCCCGCTGCCGCTGGGGGTACGCTCATCAAGATG 887  
QY 122 ctgtgagtgctgtaagatgtgcctcagcagcttggggaacatgcaaggagctgcat 181  
DB 886 CTGTGAGTGTCTTAAGATGTGGCTCAGCAGCTTGGGGACAACTGCACGGAGCTGCAT 827  
QY 182 ctgtgaccccccacccggggtcctactgtgactcagcggggagcccgaggtacgaat 241  
DB 826 CTGTGACCCCCACCCGGGGCTCTACTGTGACTACAGGGGGACCCCGAGGTACGAAT 767  
QY 242 aggaagtgtgtgcaacagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 301  
DB 766 AGAAGTGTGTGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 707  
QY 302 cggccagctccctcagcctcagcctcagcctcagcctcagcctcagcctcagcctc 361  
DB 706 CGGCCAGTCTCTCAGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 647  
QY 362 ctgtgacccacatgt 421  
DB 646 CTGCACACCACTGT 587  
QY 422 cgtgagatattcctggcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 481  
DB 586 CGTGAAGATACCTGT 527  
QY 482 acgcaagaccgaccccccgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 541  
DB 526 AGCAAGACGCAACCCCGT 467  
QY 542 gcaaggaactgcatgactcctcacaagcctctgtgtgtgtgtgtgtgtgtgtgtgt 601  
DB 466 GCACAGCAACTGT 407  
QY 602 gggggtctcagctcagctcagctcagctcagctcagctcagctcagctcagctc 661  
DB 406 GGGGCTCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAG 347  
QY 662 cctctgcaacttgcggcagctgcatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 721  
DB 346 CMTGTGAATTTGGGCGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 287  
QY 722 gt 781  
DB 286 GT 227  
QY 782 caaacgtctctacacccacagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 841  
DB 226 CACACGCTCTATACACCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 167  
QY 842 ctacaagctcaagactcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 901  
DB 166 CTCAACATCTTAAGACTATGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 107  
QY 902 ccaggtctcattgattatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 961  
DB 106 CCAAGTCTCTATGATTAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 47  
QY 962 ctgt 1007  
DB 46 CTGT 1

RESULT 5  
X76484  
ID X76484 standard; DNA: 1766 BP.  
XX  
XX X76484;  
DT 06-AUG-1999 (first entry)  
XX



OY	1263	gcccgaatccaaagccttccaaagtcagaagtcctgctgtagcttgcct---aaatccca	1319
Db	1436	gcccagactagctcgccgcccagcctggcaagttgtcagaagttgtgataggattgcc	1495
OY	1320	agaaatggaatacgaagtgagacttttaataatcaactaattctctttagatgccaaaccaca	1379
Db	1496	aggaagaagcactcagctgagagaccagatcatctgagtcctctccatgaticcaagcc-ta	1554
OY	1380	agactctttgggtccattcagaatgaatagatggaatttggacaacaatgaatactatca	1439
Db	1555	gggagctcgagatcctcttcacagacagatgagatggatgggacacagaaataagcattta	1614
OY	1440	tttggagcctccagaaggatgacttaattggtaatttgcg-gtcagcgacacaaaacta	1498
Db	1615	tttaaccttgcacaaatgactatccctgggattcttcgcttaaaaaacataccaagaatg	1674
OY	1499	tctgattcccaataatglatgacacctcaagtcatacaacatttgcacaagtgaattgaat	1558
Db	1675	ttctgttcacactgactcgtatcatcaacaagtcacacaaacttttccagtgaggaacctat	1734
OY	1559	agttgcttaatttgattt	1577
Db	1735	agttgtgcatctcgtttt	1753
RESULT	6		
ID	X76493		
AC	X76493 standard; DNA; 1202 BP.		
XX	X76493;		
XX			
DT	06-AUG-1999 (first entry)		
DE	Human WISP-1 clone 568.4A.		
XX			
XX	WNT-1 induced secreted protein; WISP-1; WISP-2; WISP-3; CTGF; tumour;		
KW	connective tissue growth factor; cancer; melanoma; arteriosclerosis;		
KW	leukaemia; lymphoid malignancy; haematopoiesis-related disorder;		
KW	tissue growth disorder; skin disorder; desmoplasia; fibrotic lesion;		
KW	kidney disorder; bone-related disorder; osteoporosis; trauma; burn;		
KW	connective tissue disorder; catabolic state; inflammation;		
XX	testicular-related disorder; angiogenesis; immunological disorder; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	W09921998-A1.		
PD	06-MAY-1999.		
XX			
PE	29-OCT-1998; 98WO-US22991.		
XX			
PR	14-APR-1998; 98US-0081695.		
PR	29-OCT-1997; 97US-0063704.		
PR	03-FEB-1998; 98US-0073612.		
XX			
PA	(GETH ) GENENTECH INC.		
PI	Botstein DA, Cohen RL, Goddard A, Gurney AL, Hsian K;		
PI	Lawrence DA, Levine AJ, Pennica D, Roy MA, Wood WI;		
XX			
DR	WPI; 1999-337420/28.		
XX			
PT	New isolated wnt-1 induced secreted polypeptides, WISP-1, 2 and 3		
PS	Claim 8; Page 186-187; 284pp; English.		
CC			
CC	The present invention describes wnt-1 induced secreted polypeptides,		
CC	WISP-1, 2 and 3. The novel WISP polypeptides, designated WISP-1, WISP-2		
CC	and WISP-3 have homology to connective tissue growth factor (CTGF).		
CC	Products from the present invention can be used to treat WISP-related		
CC	disorders such as breast, ovarian, and colon cancer or melanoma. The		
CC	products can be used to treat arteriosclerosis. The products can also be		











PD	09-DEC-1999.
XX	
XX	03-JUN-1999; 99WO-US12150.
XX	
PR	05-JUN-1998; 9805-0088320.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Young PE:
XX	
DR	WPI; 2000-147042/13.
XX	
PT	New isolated connective tissue growth factor-4, used for treating e.g.
XX	cancers -
PS	Disclosure; Page 192; 196pp; English.
XX	
CC	The invention provides an isolated human connective tissue growth factor
CC	-4 (CTGF-4) polypeptide. The CTGF-4 cDNA is deposited under ATCC No.
CC	209818. The CTGF-4 protein can be expressed by standard recombinant
CC	methodology. The polypeptide can be used for preventing, treating or
CC	ameliorating a medical condition. (See Z58613 for a detailed description
CC	of the diseases, deficiencies, disorders and infections that can be
CC	treated or prevented using the CTGF-4 polypeptide). Sequences Z58614-619
XX	represent nucleotide sequences related to CTGF-4 cDNA.
XX	
SQ	Sequence 540 BP; 145 A; 129 C; 105 G; 146 T; 15 other;

Query Match	9.7%	Score 355.8	DB 21	Length 540
Best Local Similarity	91.0%	Pred. No. 1.1e-80		
Matches 394	Conservative 0	Mismatches 37	Indels 2	Gaps 2

OY	3216	caggtgaacatctgttggattcttcctagagt-gaaacaacgcgcctcacaacataacaa	3274
Db	439	CCGtGTAAATCCtGTtNNAtTTTCTGGGNTGGAAACCtGcCTTTACAAATTTMCA	380
OY	3275	ggtctgataaacagctcactccattgaaattcagttgnaaccacaaagctagtcttcac	3334
Db	379	GGCTGGATPAAGGCTCACTCCCAATTGGAAATTCGNTGGAAACCCNAGGCTAGGTTTAC	320
OY	3335	tgaattgcatcactcaatttggnaactgnaacttagcttcaagaatcataggaagctcgg	3394
Db	319	TNGAATTGCATCTCCAAATTTGGGAAACTGATTAAGCTTTCAAAAGATCATAGGAAGCTGG	260
OY	3395	ttgtaagaacactaggatattcttcggcaatgggtgtaagaaggttggttcagaataccagct	3454
Db	259	TTTGAGAAACtAGGAGATTAATCTTGCGCAATGGGTGGAGGAAGGTGGTCANAAATMAACCCAGT	200
OY	3455	cgcgcattggttttggagaagaagacactcttaagcagagcccggaagggaagctcccaac	3514
Db	199	CGCCATTGGTTTGTAGAAACGGAACATATCTTAATGCGAGN-CGCGAGGGCCAACTCTCAAC	141
OY	3515	ccatcggttgtaagccatgtaagaagaagaatttggatccaatgttaatgaagctcttcttaag	3574
Db	140	CCATGGGTTGAAGCCATGGAGAAAGAAATTTGCAATCCATGTAAATGAAGCTTTCTTCAAG	81
OY	3575	tcaagaattccctgcaatgtgtgagcctgattcaataaanaaattaaagaataaataaataa	3634
Db	80	TCAGAAATTCCTCCATCGTGTGGCTGATTCAATAAAAATTAAGAAATTAATTAATA	21
OY	3635	tggaaaaaataaaaa 3647	
Db	20	TGGAAAAAATAAAAA 8	

RESULT 13  
258615/c  
ID 258615 standard; CDNA; 539 BP.  
XX 258615;  
XX  
OT 11-Apr-2000 (first entry)

XX	Human CTGF-4 related CDNA clone HSKXM67R.
DE	
XX	Connective tissue growth factor-4; CTGF-4; human; immune system disorder;
KW	hematopoietic disorder; autoimmune disorder; diabetes mellitus; asthma;
KW	respiratory disorder; inflammation; hyperproliferative disorder;
KW	infection; central nervous system disease; Alzheimer's disease; AIDS;
KW	food additive; ss.
XX	
OS	Homo sapiens.
XX	
PN	W09962927-A1.
XX	
PD	09-DEC-1999.
XX	
PF	03-JUN-1999; 99WO-US12150.
XX	
PR	05-JUN-1998; 98US-0088320.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Young PE;
XX	
DR	WPI; 2000-147042/33.
XX	
PT	New isolated connective tissue growth factor-4, used for treating e.g.
XX	cancers -
XX	Disclosure; Page 194; 196pp; English.

Query Match	8.6%;	Score 316.4;	DB 21;	Length 539;
Best Local Similarity	86.0%;	Pred. No. 1.1e-70;		
Matches 368;	Conservative	0;	Mismatches 53;	Indels 7; Gaps 2

OY	3220	gtgaatcatgtagatattctctgagagagaaacaactgcacatcacaattcaagcgcg	3275
Db	428	GTGGAACCCCTTTGGTTTCCGGGGGTGNAAACCCCTGCTTNCCAAT----	NCANG 374
OY	3280	gataacagctactccaattgaaattcagtagygaaccacaagagctggttcttactgaat	3335
Db	373	CTGGNNAACGCACNCCCATTTGAATCCTNGNAACCCCAAGCTTGTTCTTACTGGA	314
OY	3340	ttagcatcgaatttgggaacggaacttagctttcaagaatcagaagctctgttga	3395
Db	313	TTGCATCTCAAT--GGGAATACACTACTACTCTTCAAAATCTATAGAACTGTGGTGA	256
OY	3400	gaacatagggttt	3455
Db	255	GAAACTAGGGATTATTGTGGCAATGGGTGAGGAAGTGTCTCAAAATTAACCAAGTGGCA	196
OY	3460	ttagcttttgaagaaacggaactatctctatgacagacccggagggcgaactctcaaccctg	3515
Db	195	TTGGTTTGAAGAAAGCGAAGCTATCTTATGCAAGNACGGAGGGCAAGTCTCAGACCAAG	136
OY	3520	ggttgaaagccatgtagaagaaatttgatccaatglaatgaagctcttctaagtca	3575
Db	135	GCTTGAAGCCATGAGAGAGAAATTGGATCCAACTATATGAAGCTTTCTTAATCTAGA	76
OY	3580	attccctcgatagtgcttgccgcatctcaataaaattgaataaataaataaagtgga	3635
Db	75	ATTTCCTCGCAATGCTGTGGCTGATTCATTAATAATAATGAATTAATATTAATGGA	16

QY 3640 aaaaaaa 3647  
 DB 15 AAAAAAA 8

RESULT 14  
 258616  
 ID 258616 standard; CDNA; 311 BP.

AC 258616;  
 XX  
 XX 11-APR-2000 (first entry)  
 XX  
 DE Human CTGF-4 related CDNA clone HAPAO05R.

XX Connective tissue growth factor-4; CTGF-4; human; immune system disorder;  
 KW hematopoietic disorder; autoimmune disorder; diabetes mellitus; asthma;  
 KW respiratory disorder; inflammation; hyperproliferative disorder;  
 KW infection; central nervous system disease; Alzheimer's disease; AIDS;  
 KW food additive; ss.

XX Homo sapiens.  
 XX MO9962927-A1.  
 XX PD 09-DEC-1999.  
 XX PF 03-JUN-1999; 99MO-US12150.  
 XX PR 05-JUN-1998; 98US-0088320.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX Ruben SM, Young PE;  
 PI  
 XX MPI; 2000-147042/13.

DR  
 XX  
 PT New isolated connective tissue growth factor-4, used for treating e.g.  
 PT cancers -  
 PS Disclosure; Page 195; 196pp; English.  
 XX

CC The invention provides an isolated human connective tissue growth factor  
 CC -4 (CTGF-4) polypeptide. The CTGF-4 CDNA is deposited under ATCC No.  
 CC 209816. The CTGF-4 protein can be expressed by standard recombinant  
 CC methodology. The polypeptide can be used for preventing, treating or  
 CC ameliorating a medical condition. (See 258613 for a detailed description  
 CC of the diseases, deficiencies, disorders and infections that can be  
 CC treated or prevented using the CTGF-4 polypeptide). Sequences 258614-619  
 CC represent nucleotide sequences related to CTGF-4 CDNA.  
 CC  
 XX

SQ Sequence 311 BP; 101 A; 54 C; 58 G; 96 T; 2 other;

Query Match 8.1%; Score 295.4; DB 21; Length 311;  
 Best Local Similarity 98.7%; Pred. No. 1.7e-65;  
 Matches 307; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2833 atggaatggtgtgtatgattacacaggaattctgtgacgtgtgacagagagattcctaa 2892  
 DB 1 attgaatggtgtgtatgattacacaggaattctgtgacgtgtgacagagagattcctaa 60  
 QY 2893 cgggaaaaagactggagatacatctctactgtgacctcccaaaacctagtcagtg 2952  
 DB 61 cgggaaaaagactggagatacatctctactgtgacctcccaaaacctagtcagtg 120  
 QY 2953 caaggtatacagtggtgtctatttaattctgatacaggaagctgtgcatgtttc 3012  
 DB 121 caaggtatacagtggtgtctatttaattctgatacaggaagctgtgcatgtttc 180  
 QY 3013 ctactttatcgagagctctctctcctcaagctacagtaaaagagatttaacagtc 3072

DB 181 ctactttnttcgagagctctctctctccaaagctacatgaaatagaaatttaacagtc 240  
 QY 3073 aaatttataatagtgctcttagcaaaagagacattata-cttcaagaatgcatat 3131  
 DB 241 aaatttataatagtgctcttagcaaaagagacattataatctcaagaatgcatat 300  
 QY 3132 gtagtataca 3142  
 DB 301 gtagtataca 311

RESULT 15  
 211246  
 ID 211246 standard; DNA; 1062 BP.

AC 211246;  
 XX  
 XX 15-NOV-1999 (first entry)

DE Human connective tissue growth factor, zctgf4, generic coding sequence.  
 XX  
 KW Connective tissue growth factor; zctgf4; chromosome 6q abnormality;  
 KW diagnosis; therapy; bone marrow fibrosis; hematopoietic cell production;  
 KW hematopoietic cell differentiation; scar tissue formation; scleroderma;  
 KW cutaneous lupus erythematosus; dermatosis; end-stage kidney failure;  
 KW human; ss.

XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1.1062  
 FT CDS /\*tag= a  
 FT /product= zctgf4  
 FT

XX MO9942583-A1.  
 XX PD 26-AUG-1999.  
 XX PF 19-FEB-1999; 99MO-US03585.  
 XX PR 20-FEB-1998; 98US-0027405.  
 PA (ZYMO) ZYMOGENETICS INC.  
 XX  
 XX Jaspers SR, Sheppard PO;  
 PI  
 XX MPI; 1999-540308/45.  
 DR P-PSDB; Y34190.

DR  
 XX  
 PT New isolated polynucleotides encoding connective tissue growth  
 PT factor homologue polypeptides, used diagnostically  
 PS Claim 5; Page 105; 122pp; English.  
 XX

CC This sequence encodes the human connective tissue growth factor (zctgf4)  
 CC of the invention. The zctgf4 coding sequence may be used to produce CTGF  
 CC homologue polypeptides according to standard recombinant DNA  
 CC methodologies. The zctgf4 DNAs may also be used diagnostically as probes  
 CC to detect the presence of similar sequences in biological samples, and to  
 CC identify abnormalities or mutations within those sequences. The sequences  
 CC can be used to detect a chromosome 6q abnormality. Antagonists and  
 CC antibodies of zctgf4 can be used to treat and diagnose bone marrow  
 CC fibrosis, modulating production or differentiation of hematopoietic  
 CC cells, prevention of scar tissue formation, cutaneous lupus  
 CC erythematosus, scleroderma, dermatosis, and end-stage kidney failure.  
 CC  
 XX

SQ Sequence 1062 BP; 291 A; 114 C; 175 G; 145 T; 437 other;  
 Query Match 7.6%; Score 277.8; DB 20; Length 1062;  
 Best Local Similarity 38.7%; Pred. No. 9.2e-61;  
 Matches 380; Conservative 179; Mismatches 400; Indels 24; Gaps 3;



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2001, 07:14:19 ; Search time 1966.83 Seconds  
(without alignments)  
13032.845 Million cell updates/sec

Title: US-09-325-019-1  
Perfect score: 3658  
Sequence: 1 cgaacttaccacagctcca.....aaaaaaaaaaaaaaaaaaaa 3658

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7991742 segs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: em\_estba:\*  
42: em\_estfun:\*  
43: em\_esthum1:\*

44: em\_esthum2:\*  
45: em\_esthum3:\*  
46: em\_esthum4:\*  
47: em\_esthum5:\*  
48: em\_esthum6:\*  
49: em\_esthum7:\*  
50: em\_esthum8:\*  
51: em\_esthum9:\*  
52: em\_esthum10:\*  
53: em\_esthum11:\*  
54: em\_esthum12:\*  
55: em\_esthum13:\*  
56: em\_esthum14:\*  
57: em\_esthum15:\*  
58: em\_esthum16:\*  
59: em\_esthum17:\*  
60: em\_esthum18:\*  
61: em\_esthum19:\*  
62: em\_esthum20:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estov1:\*  
68: em\_estov2:\*  
69: em\_estp11:\*  
70: em\_estp12:\*  
71: em\_estp13:\*  
72: em\_estp14:\*  
73: em\_estp15:\*  
74: em\_estro1:\*  
75: em\_estro2:\*  
76: em\_estro3:\*  
77: em\_estro4:\*  
78: em\_estro5:\*  
79: em\_estro6:\*  
80: em\_estro7:\*  
81: em\_estro8:\*  
82: em\_estro9:\*  
83: em\_estro10:\*  
84: em\_estro11:\*  
85: em\_estro12:\*  
86: em\_estro13:\*  
87: gb\_estc41:\*  
88: gb\_estc42:\*  
89: gb\_estc43:\*  
90: gb\_estc44:\*  
91: gb\_estc45:\*  
92: gb\_estc46:\*  
93: gb\_estc47:\*  
94: gb\_estc48:\*  
95: gb\_estc49:\*  
96: gb\_estc50:\*  
97: gb\_estc51:\*  
98: gb\_estc52:\*  
99: gb\_estc53:\*  
100: gb\_estc54:\*  
101: gb\_estc55:\*  
102: gb\_estc56:\*  
103: gb\_estc57:\*  
104: gb\_estc58:\*  
105: gb\_estc59:\*  
106: gb\_estc60:\*  
107: gb\_estc61:\*  
108: gb\_estc62:\*  
109: gb\_estc63:\*  
110: gb\_estc64:\*  
111: gb\_estc65:\*  
112: gb\_estc66:\*  
113: gb\_estc67:\*  
114: gb\_estc68:\*  
115: gb\_estc69:\*  
116: gb\_estc70:\*

117: em\_estp16:\*

118: em\_estp17:\*

119: em\_estp18:\*

120: em\_estp19:\*

121: em\_estp20:\*

122: em\_estp21:\*

123: em\_estp22:\*

124: em\_estp23:\*

125: em\_estp24:\*

126: em\_estp25:\*

127: em\_estp26:\*

128: em\_estp27:\*

129: em\_estp28:\*

130: em\_estp29:\*

131: em\_estp30:\*

132: em\_estp31:\*

133: em\_estp32:\*

134: em\_estp33:\*

135: em\_estp34:\*

136: em\_estp35:\*

137: em\_estp36:\*

138: em\_estp37:\*

139: em\_estp38:\*

140: em\_estp39:\*

141: em\_estp40:\*

142: em\_estp41:\*

143: em\_estp42:\*

144: em\_estp43:\*

145: em\_estp44:\*

146: em\_estp45:\*

147: em\_estp46:\*

148: em\_estp47:\*

149: em\_estp48:\*

150: em\_estp49:\*

151: em\_estp50:\*

152: em\_estp51:\*

153: em\_estp52:\*

154: em\_estp53:\*

155: em\_estp54:\*

156: em\_estp55:\*

157: em\_estp56:\*

158: em\_estp57:\*

159: em\_estp58:\*

160: em\_estp59:\*

161: em\_estp60:\*

162: em\_estp61:\*

163: em\_estp62:\*

164: em\_estp63:\*

165: em\_estp64:\*

166: em\_estp65:\*

167: em\_estp66:\*

168: em\_estp67:\*

169: em\_estp68:\*

170: em\_estp69:\*

171: em\_estp70:\*

172: em\_estp71:\*

173: em\_estp72:\*

174: em\_estp73:\*

175: em\_estp74:\*

176: em\_estp75:\*

177: em\_estp76:\*

178: em\_estp77:\*

179: em\_estp78:\*

180: em\_estp79:\*

181: em\_estp80:\*

182: em\_estp81:\*

183: em\_estp82:\*

184: em\_estp83:\*

185: em\_estp84:\*

186: em\_estp85:\*

187: em\_estp86:\*

188: em\_estp87:\*

189: em\_estp88:\*

190: gb\_gss25:\*

191: gb\_gss26:\*

192: gb\_gss27:\*

193: gb\_gss28:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	699	19.1	729	38	AV709609	AV709609 AV709609
2	565	15.4	565	26	AI917494	AI917494 t884f07.x
3	496.4	13.6	500	23	AI683855	AI683855 tx77h09.x
4	471	12.9	471	19	AI372923	AI372923 q299a05.x
5	461	12.6	469	22	AI624151	AI624151 t928e10.x
6	442.4	12.1	444	138	BE000447	BE000447 7h29b07.x
7	419.8	11.5	428	27	AI990199	AI990199 w838c03.x
8	415.4	11.4	417	107	BE350495	BE350495 h14f01.x
9	395	10.8	406	95	AW804917	AW804917 OV4-DM09
10	384.4	10.5	386	108	BE455934	BE455934 hy08a09.x
11	348.2	9.5	381	88	AW291472	AW291472 UT-H-B12
12	338.8	9.3	503	168	AQ726685	AQ726685 HS-5411-B
13	314.4	8.6	429	145	T55016	T55016 yd42e03.x
14	300	8.2	320	27	AI971546	AI971546 wq87b12.x
15	295.4	8.1	311	5	AA316760	AA316760 EST1883
16	292.2	8.0	297	13	AA922800	AA922800 om63h03.s
17	247	6.8	389	104	BE127024	BE127024 DEBA0773
18	234.4	6.4	606	38	AV670917	AV670917 AV670917
19	234	6.4	650	176	A2217325	A2217325 RPCI-23-9
20	208	5.7	208	20	AI473336	AI473336 t883f09.x
21	183.4	5.0	197	6	AA344487	AA344487 EST50540
22	181.2	5.0	430	38	AV669292	AV669292 AV669292
23	178.2	4.9	562	22	AI588038	AI588038 mp02e12.y
24	161	4.4	161	6	AA393740	AA393740 zv64d08.x
25	135.2	3.7	1057	137	BE898435	BE898435 601681295
26	131	3.6	528	2	AA104304	AA104304 mp02e12.x
27	125	3.4	636	40	AW146261	AW146261 um39h01.x
28	124.4	3.4	679	106	BE306545	BE306545 601102787
29	123	3.4	1017	192	CNS030ER	AI255888 Tetradon
30	121.2	3.3	753	136	BE865071	BE865071 UT-M-BH2
31	118	3.2	578	37	AV600173	AV600173 AV600173
32	117.4	3.2	648	94	AW742404	AW742404 UP56f10.y
33	113.4	3.1	647	27	AI965085	AI965085 f886f04.y
34	113.4	3.1	680	38	AV702997	AV702997 AV702997
35	113	3.1	436	6	AA373233	AA373233 EST85240
36	113	3.1	843	191	CNS02CON	AI191408 Tetradon
37	111.2	3.0	505	139	BE037009	BE037009 601456766
38	108.4	3.0	624	38	AV710966	AV710966 AV710966
39	107.4	2.9	696	4	AA221075	AA221075 mw75c04.x
40	107	2.9	475	9	AA605073	AA605073 no74d01.x
41	106.8	2.9	415	104	BE166172	BE166172 MR3-HT048
42	105.4	2.9	344	104	BE127370	BE127370 DEPA1119
43	105.4	2.9	531	7	AA419858	AA419858 v48h09.x
44	105.4	2.9	806	28	AL048446	AL048446 DKE25861
45	105.4	2.9	831	135	BE742330	BE742330 601575564

## ALIGNMENTS

RESULT 1

AV709609 AV709609 729 bp mRNA

LOCUS AV709609 ABC Homo sapiens cDNA clone ADCAR05 5', mRNA sequence.

DEFINITION AV709609

ACCESSION AV709609

VERSION AV709609.1 GI:10727426

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE	1 (bases 1 to 729)
JOURNAL	Peng,Y., Song,X., Huang,Q., Huang,C., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J., Chen,Z. and Han,Z.
COMMENT	Homo sapiens cDNA ADC clones Unpublished (2000) Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@chgc.sh.cn This clone is available at CHGC in Shanghai.
FEATURES	Location/Qualifiers
SOURCE	1..729
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="ADCARB05"
	/clone_1lb="ADC"
	/tissue_type="Adrenal gland"
	/dex_stage="Adult"
	/lab_host="SOLR"
	/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT	196 a 159 c 167 g 205 t others
ORIGIN	
Query Match	19.1%; Score 699; DB 38; Length 729;
Best Local Similarity	98.4%; Pred.No.2.8e-167;
Matches 716; Conservative	0; Mismatches 11; Indels 1; Gaps 1
OY 1337	gaactttaatactactaatcttccttttagatgaccaaccacagaactcttggtccat 1396
Db 1	GACITTTATATACACTAATTCTCTTGAGACCMAACACAAGACTTGGGTCCAT 60
OY 1397	tcaagatgaatgatggaaatttgacaacatagaaataatcatatttgaggcctgccaaa 1456
Db 61	TCAGATGAAATATGGAATTTGGAACATAAGATATCTATTTTGGAGCCTGCCAAGA 120
OY 1457	ggtactgtaatggygaattctgacgtcgacgcacacaaactacctgattccaatatgt 1516
Db 121	GSTACTGTAATGGGTAATCTGACGTCANCGCACCAAAACTATCCGATTCCAAATATGT 180
OY 1517	atgcaactcaaggtcatcaaacattgcccaagtgaatgaatgttgtctaatttgatt 1576
Db 181	ATGCACCTCAAGGTCATCAACCATTTGCCAACATGATGATTAAGTGCCTTAATTGGATT 240
OY 1577	tttaatggaagaatttatccacttaacccctggagatgttgagtttagttcttcacc 1636
Db 241	TTTTATGGAAATGTGATCCATTAATCACTTGCGCAATGTTTAGGTAAAGTTCTTCACCC 300
OY 1637	ctacacatgtaagggttacagattaggtttgtcccagtcagacaataaattgataaacat 1696
Db 301	CTACACTGTGAAGGGTACAGATTAGGTTTGTCOCACGTCAGAATATAAATTGATAAACAT 360
OY 1697	tccgtgtatgvgaaaaaaccceccagtttaactccaagacagsgaagaagtcaagccatt 1756
Db 361	TCTCTTTATGAGGAAGACCCCACAGTTAATCTCCAGAACACAGGGAAGGTGAGCCATT 420
OY 1757	tcaagaagcaccaattgacatctcaactgtaataagctgcgcgactggaagagcttggcag 1816
Db 421	TCAGAAAGACCAATTGACTCTCACACTGATACAGCTGCTGACTGGCAAGGCTTTGGCGAG 480
OY 1817	ttggscagagctctcccttgaaactctccctgtccctgtctgtgttggtgtcataagaaattgt 1876
Db 481	TTGGCAGAGCTCTCTCTTAATCTTCTCCCTTGCTGCTGGGGCTTACAGAAATNGT 540
OY 1877	aagccctctggaactgacctgtctggcccccttgagaagtgtgtcccttgacaactccttact 1936

Db	541	AAGGCTGTGGACACGCGCCTGTCTTGCCCTTGAGAGTGGTTGCTTGGAAACATCTCTACT	600
QY	1937	ctacagagcctttagagagaccagctgcagacccatgcagagcccactgaaatgacccaaga	1996
Db	601	CTTACAGAGCCTTGAGAGACCCAGCTGCAGACACATGCGACAGACCACTTGATGACCAAGA	660
QY	1997	caggttca-ggttaagggtgtgtgtcacaaccaagaagtgtgtgccttgtgtgcagcccg	2055
Db	661	CAGGTTCANCGTTAGGCGGTGTGGTCAACCAAAAAGTGTGGCTTGTAGAGAGCCCTGG	720
OY	2056	ggtgacct 2063	
Db	721	GGTGACCT 728	
RESULT	2		
AI917494/c			
LOCUS	AI917494	565 bp	mRNA EST 14-DEC-1999
DEFINITION	ts84f07.x1 NCI_CGAP_GCC Homo sapiens cDNA clone IMAGE:2237989 3',		
ACCESSION	AI917494		
VERSION	AI917494		
KEYWORDS	mRNA sequence.		
SCORE	AI917494.1 GI:5637349		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP).		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: (301) 496-1550		
	Email: Robert.Strausberg@nih.gov		
	Tissue Procurement: Christopher A. Moskalkuk, M.D., Ph.D., Michael		
	R. Emmert-Buck, M.D., Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima		
	Bonaldo, Ph.D.		
	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	www-bio.lnlnl.gov/bdrip/IMAGE/IMAGE.html		
	Insert length: 914 Std Error: 0.00		
	Seq primer: -40up from Gdbco		
	High quality sequence stop: 447.		
FEATURES	Location/Qualifiers		
source	1. 565		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2237989"		
	/clone_lbp="NCI-CGAP_GCC"		
	/tissue_type="pooled germ cell tumors"		
	/lab_host="DH10B"		
	/note="Vector: pRT3D-Pac (Pharmacia) with a modified		
	polylinker; Plasmid DNA from the normalized library		
	NCI-CGAP GC4 was prepared, and as circles were made in		
	vitro. Following HAP purification, this DNA was used as		
	tracer in a subtractive hybridization reaction. The diver		
	was PCR-amplified cDNAs from a pool of 5,000 clones made		
	from the same library (clonoids 1257096-1258631,		
	146066-1470983, and 1475592-1476743). Subtraction by		
	Bento Soares and M. Fatima Bonaldo.		
BASE COUNT	* 166 a 122 c 88 g 189 t		
ORIGIN			
Query Match	15.4%;	Score 565;	DB 26;
Best Local Similarity	100.0%;	Pred. No. 3	8e-133;
Matches 565;	Conservative 0;	Mismatches 0;	Indels 0;
	Gaps 0;		



QY 3076 ttattataaagtgtctctgcaaaagagacatttaatttcaaaagaatgcatagtat 3135  
 |||||||  
 Db 565 ttttaatttaagctcttgcaaaaagacatttaatttcaaaaagacataatgtat 506  
 |||||||  
 QY 3136 gratacatatattgtgtatgctgcatgcaagaattctgtataaagaattcactcat 3195  
 |||||||  
 Db 505 gtrttacatattttgtgtatgctgcatgcaagaattctgtataaagaattcactcat 446  
 |||||||  
 QY 3196 gaattgctctctctgtcaagtcagtgatgcatatgtatgaattctgtagaagtgaaacacc 3255  
 |||||||  
 Db 445 gaattgctctctctgtcaagtcagtgatgcatatgtatgaattctgtagaagtgaaacacc 386  
 |||||||  
 QY 3256 tgcacactcaaatccaaagctcgtgtataacagctccatccatttgaaattcagtgaaac 3315  
 |||||||  
 Db 385 tgcacactcaaatccaaagctcgtgtataacagctccatccatttgaaattcagtgaaac 326  
 |||||||  
 QY 3316 ccaaaagcagagctcttctactgaattcgcacatccatttgaaagcagaaactgcctta 3375  
 |||||||  
 Db 325 ccaaaagcagagctcttctactgaattcgcacatccatttgaaagcagaaactgcctta 266  
 |||||||  
 QY 3376 aagatcataggaagctcgtgtggaagaactaggaattatctgtgcaatggtggaagaaag 3435  
 |||||||  
 Db 265 aagatcataggaagctcgtgtggaagaactaggaattatctgtgcaatggtggaagaaag 206  
 |||||||  
 QY 3436 gtggtcagaataaaccagctcgcacatgtgttttggaagaagaaactctctatgcagacc 3495  
 |||||||  
 Db 205 gtggtcagaataaaccagctcgcacatgtgttttggaagaagaaactctctatgcagacc 146  
 |||||||  
 QY 3496 cgaagagcagagctcgaaccatcgggttggaagcagaggaagaaatttgatccaatg 3555  
 |||||||  
 Db 145 cgaagagcagagctcgaaccatcgggttggaagcagaggaagaaatttgatccaatg 86  
 |||||||  
 QY 3556 taattgaagctcttctcaagtcagaattccctcgaatggtgtgctgcatccataaaaa 3615  
 |||||||  
 Db 85 taattgaagctcttctcaagtcagaattccctcgaatggtgtgctgcatccataaaaa 26  
 |||||||  
 QY 3616 ttaagaataataataatgaatgaaa 3640  
 |||||||  
 Db 25 ttaagaataataataatgaatgaaa 1  
 |||||||

RESULT 3  
 A1683855/c 500 bp mRNA EST 16-DEC-1999  
 LOCUS tx77h09.x1 NCI\_CGAP\_Utl1 Homo sapiens cDNA clone IMAGE:2275649 3,  
 DEFINITION mRNA sequence.  
 ACCESSION A1683855  
 VERSION A1683855.1 GI:4894037  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 500)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www.bio.llnl.gov/bbrp/image/image.html  
 Insert Length: 1330 Std Error: 0.00  
 Seg primer: -40UP from G1bco  
 High quality sequence stop: 403.  
 Location/Qualifiers

source 1. 500  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="2275649"  
 /clone\_id="NCI\_CGAP\_Utl1"  
 /tissue\_type="well-differentiated endometrial  
 adenocarcinoma, 7 pooled tumors"  
 /lab\_host="DH10B"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; Salt;  
 Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"  
 BASE COUNT 158 a 91 c 101 g 150 t  
 ORIGIN

Query Match 13.6%; Score 496.4; DB: 23; Length 500;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-115;  
 Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2496 agtggtagctgggaagttgaagatcagatcgaaattggaactgcttataaccat 2535  
 |||||||  
 Db 500 agtggtagctgggaagttgaagatcagatcgaaattggaactgcttataaccat 441  
 |||||||  
 QY 2556 ttcccccgtgttttaagagctccaaatggtcagaataagaaacattgcaataatgg 2615  
 |||||||  
 Db 440 ttcccccgtgttttaagagctccaaatggtcagaataagaaacattgcaataatgg 381  
 |||||||  
 QY 2616 ctgattttttatgcatcttccctctatagctcttcagctcccttcaaaagagc 2675  
 |||||||  
 Db 380 ctgattttttatgcatcttccctctatagctcttcagctcccttcaaaagagc 321  
 |||||||  
 QY 2676 agaatatcgaattctctgataattttagtgcataagcaccaaatcacatggagacaca 2735  
 |||||||  
 Db 320 aqaattatcgaattctctgataattttagtgcataagcaccaaatcacatggagacaca 261  
 |||||||  
 QY 2736 aaatccaggaatgcqgtgtagcttaaccctccctccatcagcaaccagctccatcag 2795  
 |||||||  
 Db 260 aaatccaggaatgcqgtgtagcttaaccctccctccatcagcaaccagctccatcag 201  
 |||||||  
 QY 2796 aaatttaagaatggttcgtgttaccacaaatgctgcatattgaattgtgtatgattaca 2855  
 |||||||  
 Db 200 aaatttaagaatggttcgtgttaccacaaatgctgcatattgaattgtgtatgattaca 141  
 |||||||  
 QY 2856 caggaagattctgtcagtgatgcagagaagatccctaaacgggaagaaagagatgaatcat 2915  
 |||||||  
 Db 140 caggaagattctgtcagtgatgcagagaagatccctaaacgggaagaaagagatgaatcat 81  
 |||||||  
 QY 2916 cctcctactgtgacctcccaaaacctaagtcagtgcaaggtatacagtggtcact 2975  
 |||||||  
 Db 80 cctcctactgtgacctcccaaaacctaagtcagtgcaaggtatacagtggtcact 21  
 |||||||  
 QY 2976 aaattcttgatgaatca 2993  
 |||||||  
 Db 20 aaattcttgatgaatca 3

RESULT 4  
 A1372923 471 bp mRNA EST 16-FEB-1999  
 LOCUS q299a05.NI\_Scares\_pregnant\_uterus\_NbHPU Homo sapiens cDNA clone  
 DEFINITION IMAGE:2048672 3', mRNA sequence.  
 ACCESSION A1372923  
 VERSION A1372923.1 GI:4152789  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 471)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index





QY 3218 ggtgatacatgttagattttcttgagagtgaaacacccctgcccatacaataacagc 3277  
 DB 423 GAGGAAATCATGTTAGATTCTTGAGAGTGAAAACACCTGCCATCAAAATACAGGC 364  
 QY 3278 tggatacagctactccattgaaattcaagtgaagaccagaagcttgcttactga 3337  
 DB 363 TGGATACAGCTCCTCCATTTGCAATTCAGTGAAACCCAGAGTAGTTCTTACTGA 304  
 QY 3338 attgatactcaatttgggaaactgaacttagcttcaaaagctcagaagcttggtg 3397  
 DB 303 ATTGATCATCTCAATTTGGGAAACTGAACTTAGCTTTCAAAAGATCAAGAACTGCTG 244  
 QY 3398 gagaacactagagatctctctgcaatggtgagagagagtgctcagaataacccagtcg 3457  
 DB 243 GAGAAACTAGGAGTATTTCGCGAATGGGCGAGGAGAGTGCGACAAATACCACTGCC 184  
 QY 3458 catgtgttggagaaacggaactatctatgcagagccggaggcgaagctcacaacca 3517  
 DB 183 CATTGCTTTGAGAAACGGAACATATCTTATGACAGACCCCGAGCGCAAGCTCAAAACCCA 124  
 QY 3518 tgggttgagagcctgagagagaaatttgatccaatgtaatgaactcttctaagta 3577  
 DB 123 TGGCTTGAAGCCATGAGAGAAATTTGGATCAATGTAATGAAGCTCTTTAAAGTCA 64  
 QY 3578 gaattccctgcagtgtgtgctgattcaataaataaataaataataatgag 3637  
 DB 63 GAATTCCTCGCATGTGTGTGCTGATTCATTAATAAATAATTAATAATG 4  
 QY 3638 aag 3640  
 DB 3 AAA 1

RESULT 8  
 BE350495/c 417 bp mRNA EST 18-JUL-2000  
 LOCUS ht14f01.x1 NCI\_CGAP\_K1d13 Homo sapiens cDNA clone IMAGE:3146713 3',  
 DEFINITION mRNA sequence.  
 ACCESSION BE350495  
 VERSION BE350495.1 GI:9262348  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 417)  
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 TITLE Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert\_Strausberg@nih.gov  
 Tissue Procurement: Chris Koskeluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life  
 Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The  
 I.M.A.G.E. Consortium DNA Sequencing by: Washington University  
 Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL, send email to:  
 infoimage.llnl.gov  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 414.  
 Location/Qualifiers  
 1. 417  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3146713"  
 /clone\_1db="NCI\_CGAP\_K1d13"  
 /tissue\_type="2 pooled Wilms' tumors, one primary and one  
 metastatic to brain"  
 /lab\_host="DH10B"

BASE COUNT 143 a 82 c 64 g 128 t  
 ORIGIN  
 Query Match 11.4%; Score 415.4; DB 107; Length 417;  
 Best Local Similarity 99.8%; Pred. No. 5e-95;  
 Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 /note="Organ: kidney; Vector: PCMV-SPORT6; Site\_1: SalI;  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo df.  
 Library constructed by Life Technologies."

QY 2216 acttaataaggcttaaggctgaagaggttgagctaaaggaagaggttgcttaag 2275  
 DB 417 ACTTAATAAGGCTTTAAAGCTGAAGAGTGAAGCTCAAGGAAAGGCTTGTTAAAG 358  
 QY 2276 aatataagctcttatttcttgattagaaaataaattactgtagaattctt 2335  
 DB 357 AATATCAGGCTATATTTATTTGATTAGGAAATAAATAATTACTGTAGAAATCTTTT 298  
 QY 2336 attagagccttctctgacagacatgctcgaagcttgatgataagctcagc 2395  
 DB 297 ATTAGAGGCTTTTCGTGCGACAGATTCCTCACTGCTTGCAATGATGACTCAGC 238  
 QY 2396 aatcttcagcaaatgttgagaagttcccaattatttctgcttacaatgtgaac 2455  
 DB 237 AATCTTACAGACATGTTGAGAAAGTCCCATATATTTCTGTTTACAAATGTGAAC 178  
 QY 2456 ggaagctctagaagtgagaaactcaaccagagttaccagttgctgagaaagt 2515  
 DB 177 GGAAGCTCATAGAGTGAGAAACCTCAACAGAGTACCAGTGTGACTGGGAAAGTT 118  
 QY 2516 agattcagatcgaatgtgactgctcttataaccataattcccccgttttagagc 2575  
 DB 117 AGGATTCAGATCGAAATTTGGACTGCTTTATTAACCATATTTTCCCGCTTTTAGAC 58  
 QY 2576 ttccaaatgttcagaatagagaacatgtcgaataaagcttgatttttaagtc 2632  
 DB 57 TTCCAATGTGTCAGATAGGAAACATTCATAAATGCTGATTTTATATGC 1

RESULT 9  
 AM804917/c 406 bp mRNA EST 16-MAY-2000  
 LOCUS OVA-UM0094-280300-152-h08 UM0094 Homo sapiens cDNA, mRNA sequence.  
 DEFINITION AM804917.  
 ACCESSION AM804917.1 GI:7656696  
 VERSION AM804917.1 GI:7656696  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 406)  
 REFERENCE Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zagor, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Cavalheiro, A.F., Matsukuma, A., Bais, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.C.  
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 MEDLINE 200202663  
 COMMENT Contact: Simpson A.J.C.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Paraf. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the RPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?el=tl2-OV4-UM0094-280>)

300-152-h086t3-2000-03-286t4-1)  
Seq primer: puc 18 forward  
High quality sequence stop: 406.  
Location/Qualifiers

FEATURES  
source 1. 406  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="UM0094"  
/dev\_stage="Adult"  
/note="Organ: uterus; Vector: puc18; Site: 1: Sma1; Site: 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 117 a 95 c 70 g 124 t  
ORIGIN

Query Match 10.8%; Score 395; DB 95; Length 406;  
Best Local Similarity 98.8%; Pred. No. 7.9e-90;  
Matches 398; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3131 tctatgtatcatatatttctgtatgctgtatgcaagaattctgtataagaagaattcac 3190  
DB 406 tctatgtatcatatatttctgtatgctgtatgcaagaattctgtataagaagaattcac 347  
OY 3191 tctatgtatcatatatttctgtatgctgtatgcaagaattctgtataagaagaattcac 3250  
DB 346 tctatgtatcatatatttctgtatgctgtatgcaagaattctgtataagaagaattcac 287  
OY 3251 acacctgcctacatacaattacagagctgtatgcaagaattctgtataagaagaattcac 3310  
DB 286 acacctgcctacatacaattacagagctgtatgcaagaattctgtataagaagaattcac 227  
OY 3311 gaaacccaagaagctaggtcttctactgaattctgaattcgaattctgtatgcaagaattcac 3370  
DB 226 gaaacccaagaagctaggtcttctactgaattctgaattcgaattctgtatgcaagaattcac 167  
OY 3371 ttccaagaagctaggtcttctactgaattctgaattcgaattctgtatgcaagaattcac 3430  
DB 166 ttccaagaagctaggtcttctactgaattctgaattcgaattctgtatgcaagaattcac 107  
OY 3431 ggaagctgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 3490  
DB 106 ggaagctgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 47  
OY 3491 ggaagctgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 3533  
DB 46 ggaagctgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 4

RESULT 10  
BE465934/c 386 bp mRNA EST 27-JUL-2000  
LOCUS hy08a09.x1 NCI\_CGAP\_Gc6 Homo sapiens cDNA clone IMAGE:319696 3',  
DEFINITION mRNA sequence.  
ACCESSION BE465934  
VERSION BE465934.1 GI:9511709  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 386)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/nclogap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert\_Strausberg@nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLN, send email to: info@image.lnl.gov  
Seq primer: -40UP from gibco.  
Location/Qualifiers

FEATURES  
source 1. 386  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NCI\_CGAP\_Gc6"  
/clone\_lib="NCI\_CGAP\_Gc6"  
/tissue\_type="Pooled germ cell tumors"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalised library NCI\_CGAP\_Gc4 was prepared, and ss circles were used in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1257086-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 102 a 91 c 66 g 127 t  
ORIGIN

Query Match 10.5%; Score 384.4; DB 108; Length 386;  
Best Local Similarity 99.7%; Pred. No. 3.9e-87;  
Matches 385; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3255 ctgacctcacatacaagctgtatgcaagaattctgtataagaagaattcac 3314  
DB 386 ctgacctcacatacaagctgtatgcaagaattctgtataagaagaattcac 327  
OY 3315 cccaagaagctaggtcttctactgaattctgaattcgaattctgtatgcaagaattcac 3374  
DB 326 cccaagaagctaggtcttctactgaattctgaattcgaattctgtatgcaagaattcac 267  
OY 3375 aaagatcataggaagctgttctactgaattctgaattcgaattctgtatgcaagaattcac 3434  
DB 266 aaagatcataggaagctgttctactgaattctgaattcgaattctgtatgcaagaattcac 207  
OY 3435 ggtgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 3494  
DB 206 ggtgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 147  
OY 3495 ccggaagctgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 3554  
DB 146 ccggaagctgtcagagataccagctgcgcattgttttgaaagaaaggaactatctatgca 87  
OY 3555 gtaatgaagctgttctactgaattctgtatgcaagaattctgtataagaagaattcac 3614  
DB 86 gtaatgaagctgttctactgaattctgtatgcaagaattctgtataagaagaattcac 27  
OY 3615 attaagaatataataatggaaga 3640  
DB 26 attaagaatataataatggaaga 1

RESULT 11  
AW291472/c 381 bp mRNA EST 16-JAN-2000  
LOCUS AW291472  
DEFINITION UI-R-B12-egh-d-05-0-01.s1 NCI\_CGAP\_Sub4 Homo sapiens cDNA clone  
IMAGE:2724153 3', mRNA sequence.  
ACCESSION AW291472  
VERSION AW291472.1 GI:6698035  
KEYWORDS EST.  
SOURCE human.

	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	1 (bases 1 to 381)
	AUTHORS	NCI-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
	TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
	JOURNAL	Unpublished (1997)
	COMMENT	Contact: Robert Strausberg, Ph.D. Tel.: (301) 496-1550 Email: Robert.Strausberg@nih.gov The sequence contained an oligo-dT track that was present in the strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://www.bio.lnl.gov/bdnp/image/image.html">www.bio.lnl.gov/bdnp/image/image.html</a> Seq primer: M13 Forward POLYA=yes.
FEATURES	Location/Qualifiers	
SOURCE	1..381	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2724153" /clone_lib="NCI-CGAP_Sub4" /lab_host="DH10B (Life Technologies)" /note="Vector: pYT3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The NCI CGAP Sub4 library is a subtracted library derived from the NCI-CGAP_Sub2 library which is a subtracted library derived from the NCI-CGAP_Sub1 library, which is a mixture of 21 normalized or subtracted NCI-CGAP libraries: NCI-CGAP_C04, NCI-CGAP_Pr22, NCI-CGAP_Pr28, NCI-CGAP_C010, NCI-CGAP_C016, NCI-CGAP_Rid5, NCI-CGAP_Rid12, NCI-CGAP_Rid3, NCI-CGAP_Rid11, NCI-CGAP_Lym2, NCI-CGAP_Br2, NCI-CGAP_C08, NCI-CGAP_CLL1, NCI-CGAP_Le12, NCI-CGAP_Brn23, NCI-CGAP_Lu5, NCI-CGAP_Lu24, NCI-CGAP_Lu19, NCI-CGAP_GC4, NCI-CGAP_GC6, NCI-CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below:
		NCI-CGAP_Rid3 pool 1 : LHAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonoids 1323376-1323911, 1456008-1456775, 1500552-1502855) NCI-CGAP_Rid5 pool 1 : LHAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831, 1471368-1472903, 1492104-1493255) NCI-CGAP_Lu5 pool 1 : LHAM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991, 1520904-1522439) NCI-CGAP_GC4 pool 1 : LHAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743) NCI-CGAP_Pr22 pool 1 : LHAM 2457-2459, 2758-2759, 3063-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615) NCI-CGAP_C010 pool 1 : LHAM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1144584-1145355) Subtraction was performed as previously described [Bonaldi, Lennon & Soares (1996) : Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery Genome Research 6, 791-806.] TAG_LIB=NCI-CGAP_C010 TAG_TISSUE=colon TAG_SEQ=AAACG"
BASE COUNT	97 a 104 g 103 t	
ORIGIN	TAG_SEQ=AAACG"	
Query Match	9.5%; Score 348.2; DB 88; Length 381;	
Best Local Similarity	98.9%; Pred. No. 6.7e-78;	
Matches 361; Conservative	0; Mismatches 3; Indels 1; Gaps 1;	
747	gcattcactgaactcaacctgcyggcgatgcataagcacacgtctctataaa-cccaagta	805

D8		381	GCACCAATGAACTTCACACTTGGCGGTGGATGAGCAGAGCTCTATGCAACCCCAAGTA	322
OY		806	ctgtgagcttgcacatggacaataaggctcgaaccctcccaagtcctaagactcgagct	865
D8		321	CTGTGGAGATTGCAATAGAACATAATAGTGCTGATGCCCTACAAAGCTAAGACTATGACAGT	262
OY		866	gtcttcacagtgcttcgatggcttggtcttcctccgcaggtctctatgatttaagtctg	925
D8		261	GTCCTTCCAGTGCTCTGATGGCTGCTGCTCTCTCCGCGAGCTCCTATGATTAAATGCTTG	202
OY		926	ctctctgaacctgagctgtagaaatcccacatgacatcttctgacttgagatccacc	985
D8		201	CTTCTGTAACCTGAGCTGTAGGAATCCCAATGAAATCTTTGCTACTTGGATCTTACC	142
OY		986	tgaactctcagaaatctgccaaactaggaagagcacaaatcttggctcttggagtaacca	1045
D8		141	TGACTTCTCAACAATATGTCACACTAGGACAGCACAAATCTTGCTTGGGGAGTAACCA	82
OY		1046	atgcctctgtagagcagctgaagcccttatggccataactttaccacaaatgagccttaattac	1105
D8		81	ATGGCTTGAGAGCATGATGAGCCCTTATGGCCAATATCTTTCACCAATGAGCCTTAGTTAC	22
OY		1106	cctga tttt	
D8		21	CCTGA 17.	

  

RESULT	12
LOCUS	A0726685
DEFINITION	Hs_5411_b2_H02_T7A RPCR-11 Human Male BAC Library Homo sapiens genomic clone Plate=987 Col=4 Row=P, DNA sequence.
ACCESSION	A0726685
VERSION	A0726685.1 GI:5486354
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 503) Mahairas G.G., Wallace J.C., Smith K., Swartzell S., Holzman T., Keller A., Shaker R., Furlong J., Young J., Zhao S., Adams M.D., and Hood L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL MEDLINE COMMENT	Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
CONTACT	Contact: Mahairas GG, Wallace JC, Hood L
COMMENT	High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel.: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones are derived from the human BAC library RPCR-11. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources ( <a href="http://bacpac.med.buffalo.edu/ordering_bac.htm">http://bacpac.med.buffalo.edu/ordering_bac.htm</a> ) or from Resear h Genetics (info@resgen.com). BAC end Web Server: <a href="http://www.husc.washington.edu">http://www.husc.washington.edu</a> Plate: 987, row: P column: 4 Seq primer: T7 Class: BAC ends High quality sequence stop: 503. Location/Qualifiers .1. 503 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="plate=987 col=4 Row=P" /clone.lib="RPCR-11 Human Male BAC Library" /sex="male" note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;





Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/biopl/image/image.html](http://www.bio.llnl.gov/biopl/image/image.html)  
 Seq primer: -400p from Gibco.

## FEATURES

Source

1. 320  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="2479007"  
 /clone\_lib="NCI-CGAP\_GC6"  
 /tissue\_type="Pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Plasmid DNA from the normalized library NCI-CGAP\_GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1257096-1258631, 1469064-1470983, and 1475392-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 102 a 58 c 56 g 104 t  
 ORIGIN

Query Match 8.2%; Score 300; DB 27; Length 320;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-65;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2329 tcttttatttgggscctttctgtcgcagacatctcagtgcttgcagatgtag 2388  
 Db 320 ttttttttttttgggscctttctgtcgcagacatctcagtgcttgcagatgtag 261  
 QY 2389 ctcaactgaatcttcacgaacatgttagaagttccattattctctgtcttacaat 2448  
 Db 260 ctcaactgaatcttcacgaacatgttagaagttccattattctctgtcttacaat 201  
 QY 2449 gtgaacggaagctcagtagagtgagaaacacacagagtcacccagttgtagctg 2508  
 Db 200 gtgaacggaagctcagtagagtgagaaacacacagagtcacccagttgtagctg 141  
 QY 2509 gaaggttagagttcagatcgaatgtgactgtcttataaccatatttccccctgttt 2568  
 Db 140 gaaggttagagttcagatcgaatgtgactgtcttataaccatatttccccctgttt 81  
 QY 2569 tttagagcttccaagtgtcagatatgagaacatctgcataaagtgtgctgattttta 2628  
 Db 80 tttagagcttccaagtgtcagatatgagaacatctgcataaagtgtgctgattttta 21

RESULT 15  
 AA316760 311 bp mRNA EST 19-APR-1997  
 LOCUS AA316760 Lung Homo sapiens cDNA 5' end, mRNA sequence.  
 DEFINITION  
 ACCESSION  
 VERSION  
 AA316760.1 GI:1969089  
 KEYWORDS  
 EST.  
 SOURCE  
 ORGANISM  
 human  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
 AUTHORS  
 1 (bases 1 to 311)  
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult  
 C.J., Lee,N.H., Kirschner,E.F., Weinsock,K.G., Gocayne,J.D., White  
 O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Well,C., Clayton,R.A.,

Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald  
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S., Glodde,A.,  
 Guelm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,  
 Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
 Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
 Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,  
 Small,R.V., Spriggs,T.A., Uteback,T.R., Weidman,J.F., Li,Y.,  
 Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
 Dimke,D., Feng,D.-F., Ferlie,A., Fischer,C., Hastings,G.A., He,W.W.,  
 Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,  
 Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L.,  
 Wel,Y.F., Wing,J., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon  
 M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and  
 Venter,J.C.  
 Initial assessment of human gene diversity and expression patterns  
 based upon 83 million nucleotides of cDNA sequence  
 Nature 377 (6547 Suppl.), 3-174 (1995)  
 96026280  
 JOURNAL  
 MEDLINE  
 COMMENT  
 Bioinformatics  
 Contact: Kerlavage, AR  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850 USA  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: [arkerlav@igf.org](mailto:arkerlav@igf.org)  
 For clone availability, additional sequence and expression  
 information related to this EST, please check the TIGR Human Gene  
 Index (<http://www.tigr.org/tcdb/hgi/hgi.html>)  
 Seq primer: M3 Reverse.

## FEATURES

Source

1. 311  
 /organism="Homo sapiens"  
 /db\_xref="ATCC (host):117430"  
 /db\_xref="taxon:9606"  
 /clone\_lib="Lung"  
 /dev\_stage="adult"  
 /note="Organ: Lung; Vector: pBluescript SK-; Site\_1: EcoRI  
 /site\_2: XhoI"  
 BASE COUNT 101 a 54 c 58 g 96 t 2 others  
 ORIGIN

Query Match 8.1%; Score 295.4; DB 5; Length 311;  
 Best Local Similarity 98.7%; Pred. No. 1.9e-64;  
 Matches 307; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2833 atgaagttgtgtgattatccacaggaattctgtcagtgtagagagattcctaa 2892  
 Db 1 attgaatgtgtgtgattatccacaggaattctgtcagtgtagagagattcctaa 60  
 QY 2893 cgggaaggaagatggaatatacctccttactgtgacctccccaaacctgtagctg 2952  
 Db 61 cgggaaggaagatggaatatacctccttactgtgacctccccaaacctgtagctg 120  
 QY 2953 caaggtatcagtggtgtccatataactgatatgatacagaagacgtgtgactgttc 3012  
 Db 121 caaggtatcagtggtgtccatataactgatatgatacagaagacgtgtgactgttc 180  
 QY 3013 ctactttatcgaagctctctcttccaaagctacagaagaatataatgatacagta 3072  
 Db 181 ctactttatcgaagctctctcttccaaagctacagaagaatataatgatacagta 240  
 QY 3073 aattttatataagtgcttcgaaagaagacattata-ttcaagaagaatgcatat 3131  
 Db 241 aattttatataagtgcttcgaaagaagacattataatattttcaagaagaatgcatat 300  
 QY 3132 gtaagatata 3192  
 Db 301 gtaagatata 311

Search completed: February 25, 2001, 15:20:31

Mon Feb 26 08:43:32 2001

Job time: 29172 sec

us-09-325-019-1.rst

Page 12





```

Db      2040  tgaactacagagagtgtaacgagacatgcatgahelunereunyrrrygsetcrllys 2057

RESULT      3
US-08-167-628-1
: Sequence 1, Application US/08167628
: Patent No. 5408040
:
GENERAL INFORMATION:
APPLICANT: Grotendorst, Gary R.
APPLICANT: Bradham Jr., Douglas M.
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSSEE: Spensley Horn Jubas & Lubitz
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/167,628
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/752,427
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell, Jr. Ph.D., John W.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD-1294
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-455-5100
TELEFAX: 619-455-5110
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: DB60R32
FEATURE:
NAME/KEY: CDS
LOCATION: 130..1177
US-08-167-628-1

Query Match      6.7%; Score 243.4; DB 1; Length 2075;
Best Local Similarity 56.2%; Pred. No. 5.8e-50;
Matches 523; Conservative 0; Mismatches 396; Indels 12; Gaps 3

Db      248  caccgccgctcccgctggggggtcagcctatcacagatgctgtgtagtctgtaagatgt 141
Db      248  ccccccgcgtccggcgccgctgtagcctctgctgagcggctgcggctctcccgcgctct 307
Oy      142  ggcgcacagcagcttgaggagcaactcagcagaggctgcacatctgtatgaccccaacccggggcc 201
Db      308  ggcgcacagcagcttgaggagcagcgtgtgacacgcagcagcagccctctgcagcccgacacaggcc 367
Oy      202  tctactgtgaactacagcgggagccgcgcaggtacgcaataggaagtgtgtgacaggtgtg 261
Db      368  tcttctgtgact---tcggcttcccgccgacacccgacacccgacacccgacacccgacaccc 424
Oy      262  tgggtgtgggtgctgcttccgtgagtgagtggtgcgttacaacaagcagcttctctcagacctta 321
Db      435  atgtgtgttccctgcacattcttcggtgtgtagcgtgtatccgcagccgagacgtcttccacagca 484

```

[illegible]



```

: Patent No. 5783187
: GENERAL INFORMATION:
: APPLICANT: Grotendorst, Gary R.
: APPLICANT: Bradham Jr., Douglas M.,
: TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Spensley Horn Jubas & Lubitz
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: CA
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/712,302
: FILING DATE: 11-SEP-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/386,680
: FILING DATE: 10-FEB-1995
: APPLICATION NUMBER: US/08/167,628
: FILING DATE:
: APPLICATION NUMBER: US/07/752,427
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wetherell, Jr. Ph.D., John W.
: REGISTRATION NUMBER: 31,678
: REFERENCE/DOCKET NUMBER: PD-1294
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-455-5100
: TELEFAX: 619-455-5110
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2075 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: IMMEDIATE SOURCE:
: CLONE: DB60R32
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 130..1177
: US-08-712-302-1

```



Db	485	GCTCGAAGTACCACTGCAACGCTGGACGGGCGGGGCGTGCATGACCCCTGTGCAGAA	544
Oy	379	tcgaagtcgccccccgcgcctctgctgtgtcccccaccccgcgcgctgagacataactgcgc	438
Db	545	TGGAAGTTCGCTCCGACAGCCCTGTACTGACCCCTTCCGAGAGAGGTGCATGCCCCGGGA	604
Oy	439	actcgtctgagcagtcggtatctgtgagagcagcgcacaaagagccacgcagaaactccgacc	498
Db	605	AATGCTGGAGAGATGGGTGTGTGACGACGCCAAAGACCAAAACCGTGTGGGCGTGGCC	664
Oy	499	gtgacacagagagccttcgatctgctgtgtgtgtgagtcgagacatgacagga-----act	552
Db	665	TCCGGGCTTACCGACTGGAAAGACACGCTTTGGCCGACCCCAACTATGATTAAAGCAACT	724
Oy	553	gcatagccttacaacagcccccgtggagcccttgcctccacagccctgcgcgcctgggtctcca	612
Db	725	GCTTGCTCAACACCAACAGAGTGGAGGCCCTGTTCCAAAGACTTGGAGATGGGCATCTCCA	784
Oy	613	ctcgatctccaatgtttaacgcccagctgtcgtgcctgacagagagccgcctctgcgaact	672
Db	785	CCCGGGTTACCAATGACAAACGCCCTCTCGACGCTAGAGAAAGACGACGCCCTGTGCATGG	844
Oy	673	tcggcgcatgtagtctgtgacatccatccatcttaagcagtgaaagaagtcttcgcgtcg	732
Db	845	TCAGGCTTGGGAAGCTGACTGTGAAAGAACATTTAAAGAGGCAAAAAGTGCATCCGA	904
Oy	733	tgtaacagccagagagcatccatccatccatccacactctgcgtgcctcatccacacgcctct	792
Db	905	CTCCCAAAATATCCAAAGCCTATCAAGTTTAGCTTCTGTGCTGCACACACATGAAGACAT	964
Oy	793	atcaaccccaagctactgtgagattgtcatgtgacaataagtagtctgtcatccctccaaagtc	852
Db	965	AACGAGCTAAATTTCTGTGGAGTAGTACCCACAGCCGAGATCTGCACCCCCCACAAGACCA	1024
Oy	853	agacatctgagcgtctcccttcacagctgtccatgtagtgccttgcctcccgccagagcccat	912
Db	1025	CCACCTTCGCGGTGGAGTTCAAGTGCCTTACGCGCAGGCTCATGAAGAAAGAAATGATGTT	1088
Oy	913	ggatthaatgcctgtctctgtlaaccctgcagctgttgaagaaatcccaatgcatacttctgcact	972
Db	1085	TCATCAAAACCTGTGCTGCTGCATTAACAACGTCCCGGAGACATGACATCTTGAATGCC	1144
Oy	973	tggaaatccatccctgcactctccgaattcgc	1003
Db	1145	TGTACTACGAGAGATGATCGGAACACTGGC	1175

7  
 RESULT 08-880-031-1  
 1, Application US/08880031  
 Sequence No. 5816756  
 Patent No. 5816756  
 GENERAL INFORMATION:  
 APPLICANT: Grotenordst, Gary R.  
 APPLICANT: Brahman Jr., Douglas M.  
 TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Spensley Horn Jubes & Lublitz  
 STREET: 4225 Executive Square, Suite 1400  
 CITY: LA Jolla  
 STATE: CA  
 COUNTRY: US  
 ZIP: 92037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/880,031  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:

```

1      APPLICATION NUMBER:  US/08/167,628
2      FILING DATE:
3      ATTORNEY/AGENT INFORMATION:
4      NAME:  Wetherell, Jr. Ph.D., John W.
5      REGISTRATION NUMBER:  31,678
6      REFERENCE/DOCKET NUMBER:  PD-1394
7      TELECOMMUNICATION INFORMATION:
8      TELEPHONE:  619-455-5100
9      TELEFAX:  619-455-5110
10     INFORMATION FOR SEQ ID NO. 1:
11     SEQUENCE CHARACTERISTICS:
12     LENGTH:  2075 base pairs
13     TYPE:  nucleic acid
14     STRANDEDNESS:  single
15     TOPOLOGY:  linear
16     MOLECULE TYPE:  CDNA
17     IMMEDIATE SOURCE:
18     CLONE:  DB60Rj2
19     FEATURE:
20     NAME/KEY:  CDS
21     LOCATION:  130..1177
22
23     US-08-880-031-1

```

[illegible]

Db 905 CTCGCAAAATCTCAAGCCTATCAAGTTGAGCTTCTGCGTCACACGACATGAAGACAT 964  
 Qy 793 atcaacccaagactggtgagttgcatgagacaatagtgctgcatccctcaagctca 852  
 Db 965 ACCGAGCTAAATCTGTGAGATGATACGACGGCGGATGTCACACCCCAAGAACCA 1024  
 Qy 853 agactatcagctgtctccctcagctgctgcatgagcttgcttcccgccagctcat 912  
 Db 1025 CCACCTGCGGGTGGAGTTCAAGTCCCTGACGGCGAGGTCAATGAAGAATCATGATGT 1084  
 Qy 913 ggattaatgctgtctctgtaacatgagctgtagaatacccaatgacatcttgctgact 972  
 Db 1085 TCATCAAGACCTGTGCTGCTCCATTTACACTGTCCGAGAGACATGACATCTTTGAATCGC 1144  
 Qy 973 tggaaatccactgactgactctcagaatgctc 1003  
 Db 1145 TGTACTACGAAGATGTACGAGACATGCG 1175

## RESULT 8

US-09-097-179-1

Sequence 1, Application US/09097179

Patent No. 6149916

GENERAL INFORMATION:

APPLICANT: Grotendorst, Gary R.

APPLICANT: Bradham Jr., Douglas M.

TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Spensley Horn Judas & Lubitz

STREET: 4225 Executive Square, Suite 1400

CITY: La Jolla

STATE: CA

COUNTRY: US

ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/097,179

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/386,680

FILING DATE: 10-FEB-1995

APPLICATION NUMBER: US/08/167,628

FILING DATE:

APPLICATION NUMBER: US/07/752,427

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Wetherell, Jr. Ph.D., John W.

REGISTRATION NUMBER: 31,678

REFERENCE/DOCKET NUMBER: PD-1294

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-455-5100

TELEFAX: 619-455-5100

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2075 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

IMMEDIATE SOURCE:

CLONE: DB60R32

FEATURE:

NAME/KEY: CDS

LOCATION: 130..1177

US-09-097-179-1

Query Match 6.7%; Score 243.4; DB 3; Length 2075;  
 Best Local Similarity 56.2%; Pred. No. 5.8e-50;  
 Matches 523; Conservative 0; Mismatches 396; Indels 12; Gaps 3;

Qy 82 caaccgcctgcgcctggtggtgagctcctacacagatgctgtagtgcgtgtaagatgt 141  
 Db 248 CCCCCCGTCCCGCGGGCGGCGGAGCCCTGCTGAGAGCGCGGCTGCTGCTGCTGCTGCT 307  
 Qy 142 ggcctcagcagcttggtggtggtggtggtggtggtggtggtggtggtggtggtggt 201  
 Db 308 GCGCCAGAGCTGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367  
 Qy 202 tctactgtaactacagcgggagcggcggcggcggcggcggcggcggcggcggcggcggc 261  
 Db 368 TCTTCTGACT---TCGGCTCCCGCGGCGGCGGAGCCGAGATCGGCTGCTGCTGCTGCT 424  
 Qy 262 tcggtgtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 321  
 Db 425 ATGGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484  
 Qy 322 actgcaagtacacatgacatgacatgacatgacatgacatgacatgacatgacatgacat 378  
 Db 485 GCTGCAAGTACCAATGCAAGTACCAATGCAAGTACCAATGCAAGTACCAATGCAAGTACCA 544  
 Qy 379 tcgagtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 438  
 Db 545 TGGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604  
 Qy 439 actgctgtgagcagctggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 498  
 Db 605 AATGCTGGAG 664  
 Qy 499 gtgacacag 552  
 Db 665 TCGCGGCTTACGAGCTGAG 724  
 Qy 553 gcatagcctacacag 612  
 Db 725 GCT 784  
 Qy 613 ctccgagctcctcaatgttaacgcccagctgctgctgctgctgctgctgctgctgctgctgct 672  
 Db 785 CCCGGTTACCAATGCAAG 844  
 Qy 673 tggagcagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 732  
 Db 845 TCAGGCTTGGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCAAGCTGCA 904  
 Qy 733 tgtacagcagagagatcctacag 792  
 Db 905 CTCGCAAAATCTCAAGCCTATCAAGTTGAGCTTCTGCGTCACACGACATGAAGACAT 964  
 Qy 793 atcaacccaagactggtgagttgcatgagacaatagtgctgcatccctcaagctca 852  
 Db 965 ACCGAGCTAAATCTGTGAGATGATACGACGGCGGATGTCACACCCCAAGAACCA 1024  
 Qy 853 agactatcagctgtctccctcagctgctgcatgagcttgcttcccgccagctcat 912  
 Db 1025 CCACCTGCGGGTGGAGTTCAAGTCCCTGACGGCGAGGTCAATGAAGAATCATGATGT 1084  
 Qy 913 ggattaatgctgtctctgtaacatgagctgtagaatacccaatgacatcttgctgact 972  
 Db 1085 TCATCAAGACCTGTGCTGCTCCATTTACACTGTCCGAGAGACATGACATCTTTGAATCGC 1144  
 Qy 973 tggaaatccactgactgactctcagaatgctc 1003  
 Db 1145 TGTACTACGAAGATGTACGAGACATGCG 1175

## RESULT 9

PCT-US96-08140-1

Sequence 1, Application PC/TUS9608140

```

GENERAL INFORMATION:
APPLICANT: University of South Florida
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: FISH & RICHARDSON P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08140
FILING DATE: 30-MAY-1996
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Hall, Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07414/003WO1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-678-5070
TELEFAX: 619-678-5070
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2075 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
CLONE: CTGF
FEATURE:
NAME/KEY: CDS
LOCATION: 130..1177
PCT-US96-08140-1

Query Match 6.7%; Score 243.4; DB 4; Length 2075;
Best Local Similarity 56.2%; Pred. No. 5,8e-50;
Matches 523; Conservative 0; Mismatches 396; Indels 12; Gaps 3;

```

```

499 gtgacacaggagcctcgatgctggtggtgaggcaltgacacagga-----act 552
665 tcggcgcttaccgactggaagacagcttgcccgacacacactatgattagaccact 724
553 gcatggtctacacagcccttgaggccttgctccacacagcggcgccgtggtctcca 612
725 gccctgctccagacacagatgagtgagcctgttccaaagactgtgagatggcatctcca 784
613 ctggatctccaatgttaagcccgagctgctggccttgagcaagagacgacctgcaact 672
785 cccgggttaccatgacacagcctctctgcagcctgagaaagacagccctgtgcatgg 844
673 tccgacatgagatgtgagcatccatcacatcaatgaaggcaggaaagatgctggctg 732
845 tcagccttgccagagctgacatcggaagaaacatttmaagagggcaaaaagtcattccgta 904
733 tttacacagcagagatccatcagaaattcaactctggcggtcgatccagcaagcctct 792
905 ctcccaaaattcccaagcctatccaaattgagcttctgacacacacacacacacat 964
793 atcaacccaagatctgtgagatgttgatgacatagatgctgcatccctacaaagtcta 852
965 accgacctaaattctgtgagatgtagacagcagccgagctgacaccccccacagacaa 1024
853 agactatcgagtgctctccagtgctcgtatggcttgcttccgcaggtctcat 912
1025 ccacctgcgggtgagtgatcaagtgccctgacgagcgagctcatgaagaacacatgatgt 1084
913 ggtatattgcctgctctgttaacctgagtgtagaattccaaatgacatcttgctgact 972
1085 tcattcaagacctgtgctcccatctacactgtcccgagacaaatgacattttgaatccg 1144
973 tggatccctacacctgactctccagaaattgc 1003
1145 tttactacagaaagatgtacgagacatggc 1175

RESULT 10
US-08-459-101A-1
Sequence 1, Application, 'us/08459101A
Patent No. 5945300.
GENERAL INFORMATION:
APPLICANT: LI, ET AL.
TITLE OF INVENTION: Connective Tissue Growth Factor-2
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: CAPILLA, BYRNE, BAIN, GILFILLAN,
ADDRESSER: CROCHT, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,101A
FILING DATE: June 2, 1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07736
FILING DATE: 12 JUL 94
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-317
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

```

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1128 BASE PAIRS  
TYPE: NUCLEIC ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: LINEAR  
MOLECULE TYPE: CDNA  
US-08-459-101A-1

Query Match 2.8%; Score 102.2; DB 2; Length 1128;  
Best Local Similarity 53.0%; Pred. No. 9.9e-16;  
Matches 218; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

QY 548 gaactgatatgcttaccacgaagcccttgagcccttgctccacagctgagcgctgggt 607  
DB 684 GAATGATTGTTCAACAACTTCATGCTCCAGTCTCAAGACCTGTGGAACCTGGTAT 743  
QY 608 ctccactgcatctccacatgttaacgcagctgctgagcgaagagagccgctctg 667  
DB 744 CTCACACGAGTACCAATACAAACCTGAGTGCCTGCTGTGAAGAAACCGGATTTG 803  
QY 668 caactgagcgcatgagatgtgacatcatalcaatcaatgaagcaggaagaagtgtc 727  
DB 804 TGAGTGCCTGTTGTGAGACGACGAGTACAGCAGCTGAAAAAGGGCAAGAAATGAC 863  
QY 728 gactgtgtaccagccagagcatcaccatcaccatcgagcgctgcatcagcaagc 767  
DB 864 CAGACCAAAATCCCCCAACACGACGAGTCTTACTACGCTGAGTGTGAGTGA 923  
QY 788 ctccatcaacacgaagctgtgagtttgatgacacatagtgatcctccatca 847  
DB 924 GAATACGGGCCCAAGTACGCTGCTGCTGAGACGGCGAGTCTGCAACGCCACGT 983  
QY 848 gtctaagactatgagcgtgctcctccagttccctgattggctgctcccgccaggt 907  
DB 984 GACCAGCAGCTGTGAAGATGGGTTCCCTCGAAGATGGGAGACATTTTCAAGAACGT 1043  
QY 908 ccatgattaatgagcgtctgttaaccctgagcagtgatgaagatcccatga 958  
DB 1044 CATGATGATCAGTCTTCAAAATGCACTACACTGCCCGCATGCCAARGA 1094

RESULT 11  
US-08-656-393-1  
Sequence 1, Application US/08656393

PATENT No. 5837258  
GENERAL INFORMATION:  
APPLICANT: Grotendorst, Gary R.  
TITLE OF INVENTION: METHODS AND USES OF CONNECTIVE TISSUE  
TITLE OF INVENTION: GROWTH FACTOR AS AN INDUCTION AGENT  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/656,393  
FILING DATE: 31-MAY-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Halluin, Albert P.  
REGISTRATION NUMBER: 25,227  
REFERENCE/DOCKET NUMBER: 8389-040  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-854-3660  
TELEFAX: 415-854-3694  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2541 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-656-393-1

Query Match 1.6%; Score 57.8; DB 2; Length 2541;  
Best Local Similarity 68.6%; Pred. No. 8.9e-05;  
Matches 94; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 82 caaccgctgcccgcctgggggtcagctcatcacagatgctgtgagctgtaaatgt 141  
DB 2101 CGCGGCGCTGCCCGGGGGGGGTGAGCTGCTGTGAGCGGCTGCGCTGCGCGGTCT 2160  
QY 142 ggcctagcagcttgaggacactgcagcagagctgcatctgtgacccacagggcc 201  
DB 2161 GCGC-AAGCAGCTGGCGAGCTGTGACCGAGCGACCCCTGCGACCGACAAAGGCC 2219  
QY 202 tctactgtacacagc 218  
DB 2220 TCTTCTGTGACTTGGC 2236

RESULT 12

US-08-117-373-7/C  
Sequence 7, Application US/08117373

PATENT No. H001903  
GENERAL INFORMATION:  
APPLICANT: BOWDEN, PAUL E.  
APPLICANT: LUC, XIAOCHUN  
APPLICANT: MAWZYNYAK, CYNTHIA J.  
TITLE OF INVENTION: ISOLATED HAIR KERATIN GENES AND THEIR  
TITLE OF INVENTION: USE IN HAIR GROWTH ACTIVE IDENTIFICATION ASSAYS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE PROCTER & GAMBLE COMPANY  
STREET: 11810 EAST MIAMI RIVER ROAD  
CITY: CINCINNATI  
STATE: OHIO  
COUNTRY: USA  
ZIP: 45239-8707  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/117,373  
FILING DATE:

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: CONSTANDE, BRAHM J.  
REGISTRATION NUMBER: 34,804  
REFERENCE/DOCKET NUMBER: 4518R26  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (513) 627-2858  
TELEFAX: (513) 627-0260  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 291 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-117-373-7

Query Match	1.4%	Score 52	DB 1	Length 291
Best Local Similarity	64.0%	Pred. NO. 0.00088		
Matches	110	Conservative	0	Mismatches 60; Indels 2; Gaps 2
Oy	2351	tggtccacgaacattgctctcagctcagcttgatgatattatagctcactgaatcttcacgacaat	2410	
Db	170	TTGTGCAGGCTTAGTTCCTCAGTAGATTACATGGAATTAACCTTTTAATCTTTACAACTTT	111	
Oy	2411	gttggaagatcccatattatctctgtcctacaaatgtgaaacggaaagctcatagag	2470	
Db	110	ATGAGCTGGGTCTATTATTATCT-TCATTTTGAAGAAATATATAAATGAGTACATGAGG	52	
Oy	2471	tgagaaacactcaacagagctcacccagcttgctgactgggaaagttgagcttc	2522	
Db	51	CTAAGTAATTGCTCTGAG-TTTACACAGCTAGAGAGTGGCAGACGCTTGAATTC	1	

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match	1.3%	Score 48.2	DB 1	Length 7218
	Best Local Similarity	2.5%	Pred. No. 0.03		
	Matches	8	Conservative 191	Mismatches 124	Indels 0
OY	3356 gatttcgcaccccaattgggaaccggaacttagcttcaaagatcatatgaagaattcgtgt	3395			
	'                 : : : : : : : : : : : : : : : : : :				

[illegible]

```

RESULT 14
US-09-009-913-1
Sequence 1, Application US/09009913
Patent No. 608/485
GENERAL INFORMATION:
APPLICANT: Axy's Pharmaceuticals, Inc.
TITLE OF INVENTION: Asthma Related Genes
NUMBER OF SEQUENCES: 339
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bozicevic & Reed, LLP
STREET: 285 Hamilton Ave, Suite 200
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,913
FILING DATE: 21-JAN-1998
CLASSIFICATION: ''
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sherwood, Pamela J
REGISTRATION NUMBER: 36,677
REFERENCE/DOCKET NUMBER: SEQ-4P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-327-3231
TELEFAX: 650-327-3231
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7298 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-09-009-913-1

```

Query Match	1.3%	Score 46;	DB 3;	Length 7928;
Best Local Similarity	52.9%	Pred. No. 0.29;		
Matches 145;	Conservative	0;	Mismatches 125;	Indels 4;
				Gaps 2.

OY 2228 tttaaacgtcgaagaggttcgaagcctaagaaggaaggtctcttctaatacgaatacagccta 2287

Db 27223 TTTAAAGTATGCTTATTAATCTGCAAAACAGCTCATAGAGCTTTTATTTATTT 27282  
QY 2288 tttttattgtattgagaaataataattactgttagaattctttatttgaggccct 2347  
Db 27283 TATTTTATTTTGTGCTTTTAAACAGATATATTAACAAACATTTGTTTGTGCTTC 27342  
QY 2348 ttttgtgcagacattgctctcagtgcttgcatgtattagctcactgaattctcaagac 2407  
Db 27343 CTGTGACACAGGCTCTGTGTTAAGCACTTAACAT-CACATATATCAAGCACTTTGCTAAT 27401  
QY 2408 aatgttgagaaagttcccatattattctgttcttacaatgagaaagcagctcag 2467  
Db 27402 AAGAGCTGTAATAGTATTAATTAATTTCTGCTTT--ACAGCTGCAACAGAGACTCAGAG 27458  
QY 2468 aggtgagaaactcaaccagagtcaccagcttg 2501  
Db 27459 AGGTTAGTAACCTTGCCCAAGTCAACAGACTG 27492

## RESULT 15

US-08-232-463-14  
Sequence 14, Application US/08232463  
Patent No. 5670367  
GENERAL INFORMATION:  
APPLICANT: DORNER, F.  
APPLICANT: SCHEIFLINGER, F.  
APPLICANT: FALKNER, F. G.  
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,463  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,313  
FILING DATE:  
APPLICATION NUMBER: EP 91 114 300.6  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 30472/114 IMM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)836-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
CLONE: PTZgpt-F1s  
US-08-232-463-14

Query Match 1.2%; Score 45.2; DB 1; Length 7218;  
Best Local Similarity 8.5%; Pred. No. 0.16;  
Matches 38; Conservative 21; Mismatches 199; Indels 0; Gaps 0;

QY 885 gggcttgctctcccgccaggtccatgatgaatgaatgctgctctgtaaccagctgt 944  
Db 1035 GAGCTTGCTGCGAGGTCGAGGAGCTTGCGATTTTTTTTTTTTTTTTTTTTTT 1094  
QY 945 aggaatcccaatgacatctctgtcgtcgttgaaatccatccctgactctcagaatggc 1004  
Db 1095 YY 1154  
QY 1005 aactagcgagcaaatcttggtctctgggactaaccatgacctgtgaagcagtcag 1064  
Db 1155 YY 1214  
QY 1065 cccctatggcgaataactcttccacaaatgagccctagtaaccctgactgagaccctggc 1124  
Db 1215 YY 1274  
QY 1125 ctccattctgtctgaacatcaatgaagcctgagtgtcgtcagggccatgcta 1184  
Db 1275 YY 1334  
QY 1185 tgaatttctccttgatcatcactcagcatctactcctaagaaatgacctgtctcagct 1244  
Db 1335 YY 1394  
QY 1245 gtctggaactacacccagcctgactcagccttccaaagtcactagaagtcctcgtcgtgat 1304  
Db 1395 YY 1454  
QY 1305 ctgcttaaatcccaagaatgaaatca 1332  
Db 1455 TTTACTACTTGATAGTATGTAATTA 1482

Search completed: February 25, 2001, 16:36:27  
Job time: 23368 sec

